(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
14 October 2004 (14.10.2004)

PCT

(10) International Publication Number
WO 2004/087892 A1

(51) International Patent Classification[7]: C12N 1/20

(21) International Application Number: PCT/KR2004/000752

(22) International Filing Date: 31 March 2004 (31.03.2004)

(25) Filing Language: Korean

(26) Publication Language: English

(30) Priority Data:
10-2003-0020022 31 March 2003 (31.03.2003) KR
10-2003-0020023 31 March 2003 (31.03.2003) KR

(71) Applicant *(for all designated States except US)*: **ALGENTECH** [KR/KR]; 599-1 Daeyeon-dong, Nam-gu, Busan 608-020 (KR).

(71) Applicants and
(72) Inventors: **KIM, Young Tae** [KR/KR]; #107-1002, Kyungnamsunkyung Apt, Jwa-dong, Haeundae-gu, Busan 612-030 (KR). **LEE, Jae Hyung** [KR/KR]; 167-2 Sambang-dong, Gimhae-si, Gyeongsangnam-do 621-190 (KR).

(74) Agent: **LEE, Won-Hee**; 8th Fl. Sung-ji Heights II, 642-16 Yoksam-dong, Kangnam-ku, Seoul 135-080 (KR).

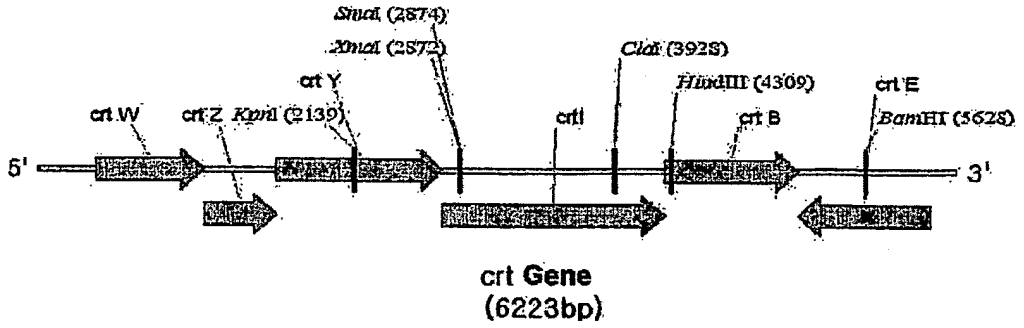
(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

**Published:**
— *with international search report*

*For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.*

(54) Title: GENE INVOLVED IN THE BIOSYNTHESIS OF CAROTENOID AND MARINE MICROORGANISM, PARACOCCUS HAEUNDAENSIS, PRODUCING THE CAROTENOID




crt Gene
(6223bp)

(57) Abstract: The present invention relates to a gene involved in the biosynthesis of carotenoid and a marine microorganism producing the carotenoid, more particularly to a gene represented by Seq ID. No. 5, 7, 9, 11, 13 and 15 which encoding the protein needed to biosynthesis of carotenoid and a marine microorganism, Paracoccus haeundaensis, producing the carotenoid. Since the gene and the microorganism can effectively be used for, the massive-production of carotenoid.

WO 2004/087892 A1

# GENE INVOLVED IN THE BIOSYNTHESIS OF CAROTENOID AND MARINE MICROORGANISM, PARACOCCUS HAEUNDAENSIS, PRODUCING THE CAROTENOID

## FIELD OF THE INVENTION

The present invention relates to genes involved in the biosynthesis of carotenoid and a marine microorganism producing the carotenoid.

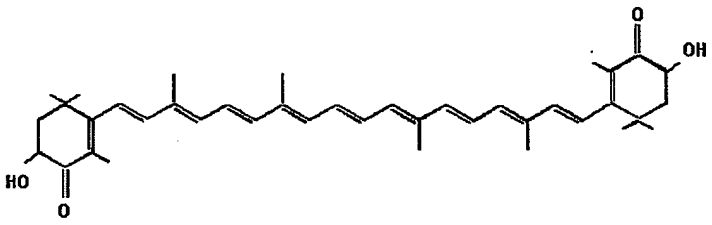
## BACKGROUND

Carotenoid, a C40 isoprenoid compound having an anti-oxidant activity, means a group of pigment that is widely distributed in the nature. More than 600 kinds of carotenoids have been known so far, and they are all in different forms. The color of carotenoid varies from its molecular structure; that is whether it is yellow, red, scarlet or orange is decided upon the molecular structure of carotenoid. The examples of carotenoids are β-carotene (an orange pigment included in a carrot), licopene (a red pigment included in a tomato), fucoxanthin (a yellowish brown or a brown pigment included in marine

plants), etc. As a precursor of vitamin A in human, carotenoid has activities of preventing oxidation, scavenging harmful oxygen, inhibiting the proliferation of cancer cells, and preventing the development of a cancer. It suggests that it has preventive effect on cardiovascular diseases, cancers and other adult diseases. It has been disclosed recently that carotenoid enhances immunity as being exposed on UV, so that it reduces skin damages by UV or inhibits the production of melanin. Since then, carotenoid came into a spotlight as a cosmetic material in Europe and in the U.S.A. Carotenoid is now in use as a health food ingredient (nutritional supplement), a pharmaceutical composition and a food-coloring agent, or as a pigment for animal feeds.

Among many carotenoids, astaxanthine (3, 3'-dihydroxy-β ,β -carotene-4, 4'-dione) having the structure of below <Chemical Formula 1> is a scarlet or light orange color pigment produced in nature.

<Chemical Formula 1>




Astaxanthine is mostly included in tissues of marine animals such as shrimps, red seabreams, salmons and lobsters, etc (Fujita *et al.*, *Nippon Suisan Gakkaishi.*, 49: 1855-1869, 1983; Johnson, E. A., *Crit. Rev. Biotechnol.*, 11: 297-326, 1991; Nelis *et al.*, *J. Appl. Bacteriol.*, 70: 181-191, 1991). Astaxanthine not only inhibits the reactions of active oxygen to destroy DNA, proteins and lipids in cells during aerobic metabolism, to cause aging in cells and tissues, and to induce a cancer but also suppresses the generation of hydroxy or peroxy radicals (hydroxy or peroxy radicals (Palozza *et al.*, *Arch. Biochem. Biophys.*, 297: 291-295, 1992; Shimidzu *et al.*, *Fish Sci.*, 62: 134-137, 1996). In addition, astaxanthine has been known to have immune modulatory activity and cardioprotective effect (Jyonuchi *et al.*, *Nutr. Cancer.*, 19: 269-280, 1993). In particular, an antioxidant activity of

astaxanthine is 10 times as high as that of other carotenoids and 100 times as high as that of α-tocopherol. However, toxicity of astaxanthine has not been reported as of today. Astaxanthine has been widely used for the treatment and the prevention of various diseases including neurodegenerative diseases, cancers, immune disorders, cardiovascular diseases, etc, and studies are still going on further (Beal, H. F., *The Neuroscientist*, 3: 21-27, 1991; Chew *et al.*, *Anticancer Res.*, 19: 1849-1853, 1999; Murillo E., *Arch. Latinoam. Nutr.*, 42: 409-413, 1992). Astaxanthine is also being industrially used as a coloring agent and has been registered as a food additive in the name of 'Phaffia color' in Korea. The consumption of astaxanthine increases over 15% every year, suggesting the importance of astaxanthine.

A method for chemical synthesis of astaxanthine has been recently developed by a company (Hoffman-LaRoche, Switzerland). However, a synthesized astaxanthine showed lower *in vivo* absorption and weaker stability as a food additive than a natural astaxanthine, so that the use of the synthesized astaxanthine was allowed just in

some of European countries. Thus, a way to synthesize a natural astaxanthine is in strong demand and especially a way to produce astaxanthine using a microorganism producing astaxanthine becomes the focus of industrial interest. *Phaffia rhodoxyma* (Miller *et al.*, *Int. J. Syst. Bacteriol.*, 48: 529-536, 1976), a kind of yeast, *Haematococcus pluvialis* (Bubrick, *Bioresour Technol.*, 38: 237-239, 1991), a kind of *Chlorophyta*, Gram-positive *Brevibacterium* 103 (Lizuka & Nishimura, *J. Gen. Appl. Microbiol.*, 15: 127-134, 1969), Gram-negative *Agrobacterium aurantiacum* (Yokoyama *et al.*, *Biosci. Biotechnol. Biochem.*, 58: 1842-1844, 1994), *Paracoccus marcusii* (Harker *et al.*, *Int. J. Syst. Bacteriol.*, 48: 543-548, 1998) and *Paracoccus carotinifaciens* (Tsubokura *et al.*, *Int. J. Syst. Bacteriol.*, 49: 277-282, 1999) are the examples of the microorganisms producing astaxanthine.

Studies have been focused on a gene coding an enzyme involved in biosynthesis of carotenoid for the past 6 years. As a result, a number of genes involved in biosynthesis of carotenoid were cloned from various microorganisms and functions of them

were also examined (Armstrong, G. A., *J. Bacteriol.*, 176: 4795-4802, 1994; Sandmann, G., *Eur. J. Biochem.*, 223: 7-24, 1994; Wieland, B., *J. Bacteriol.*, 176: 7719-7726, 1994). The pathway of carotenoid biosynthesis is derived from FPP (farnesyl pyrophosphate), which is an intermediate product of general isoprenoid synthesis pathway. As seen in FIG. 8, FPP and IPP (isopentenyl pyrophosphate) turn into GGPP (geranylgeranyl pyrophosphate) by geranylgeranyl pyrophosphate synthase encoded by *crtE*. Then, GGPP turns into β -carotene by the reactions of phytoene synthase encoded by *crtB*, phytoene desaturase encoded by *crtI* and lycopene cyclase encoded by *crtY*. β -carotene changes into astaxanthine finally by the reactions of β -carotene ketolase encoded by *crtW* and β -carotene hydroxylase encoded by *crtZ*.

Nucleotide sequences, an organization and characteristics of *crt* gene (carotenogenic gene) involved in biosynthesis of carotenoid have been investigated in *Rhodobacter capsulatus* (Armstrong *et al.*, *Mol. Gen. Genet.*, 216: 254-268, 1989), *Erwinia herbicola* (Sandimann *et al.*, *FEMS Microbiol. Lett.*, 71: 77-82, 1990; Hundle *et al.*, *Photochem. Photobiol.*, 54: 89-93, 1991) and

*Ervinia uredovora* (Misawa *et al.*, *J. Bacteriol.*, 172: 6704-6712, 1990). Besides, *crt* gene involved in biosynthesis of carotenoid, which is composed of *crtB*, *crtI*, *crtY*, *crtW* and *crtZ*, has been isolated from *Agrobacterium aurantiacum*, a marine microorganism (Norihiko *et al.*, *J. Bacteriol.*, 177(22): 6575-6584, 1995). Another report has been made on three genes (*crtB*, *crtI* and *crtY*) coding enzymes catalyzing reactions from GGPP to β-carotene and *Phaffia rhodozyma* where those three genes are inserted (WO 97/23633).

Thus, the present inventors isolated and identified a novel *Paracoccus* genus microorganism producing astaxanthine and studied further to separate a gene involved in biosynthesis of carotenoid from the microorganism. As a result, the present inventors successfully cloned *crtE*, *crtB*, *crtI*, *crtY*, *crtW* and *crtZ* genes and *crt* gene containing all of the above genes as well, and then the inventors examined nucleotide sequences of them, too. The present inventors completed this invention by confirming that carotenoid could be produced by using the *crt* gene in microorganisms that were not able to produce

carotenoid.

## SUMMARY OF THE INVENTION

It is an object of this invention to provide a *Paracoccus haeundaensis* producing astaxanthine.

It is also an object of this invention to provide a protein needed in biosynthesis of carotenoid and a gene having nucleotide sequences selected from a group consisting of nucleotide sequences represented each by SEQ. ID. No 5, No 7, No 9, No 11, No 13 and No 15.

It is a further object of this invention to provide a gene involved in biosynthesis of carotenoid containing the above gene. In particular, the present invention provides a *crt* gene having nucleotide sequences represented by SEQ. ID. No 4.

It is another object of this invention to provide a recombinant vector containing the gene involved in biosynthesis of carotenoid above.

It is another object of this invention to provide a method for producing carotenoid using the above gene involved in biosynthesis of carotenoid.

## DETAILED DESCRIPTION OF PREFERRED EMBODIMENTS

In order to achieve the above object, the present invention provides a *Paracoccus haeundaensis* producing astaxanthine.

The present invention also provides a protein needed in biosynthesis of carotenoid and a gene having nucleotide sequences selected from a group consisting of nucleotide sequences each represented by SEQ. ID. No 5, No 7, No 9, No 11, No 13 and No 15.

The present invention further provides a gene involved in biosynthesis of carotenoid containing the above gene. In particular, the present invention provides *crt* gene having nucleotide sequences represented by SEQ. ID. No 4.

The present invention also provides a recombinant vector containing the gene involved in biosynthesis of carotenoid above.

The present invention further provides a method for producing carotenoid using the above gene involved in biosynthesis of carotenoid.

Hereinafter, the present invention is

described in detail.

The present invention provides a *Paracoccus haeundaensis* producing astaxanthine.

A microorganism was isolated from seawater sample taken from Haeundae shore in Busan, Korea as a strain having orange or red color. After investigating characteristics of the strain, it was confirmed that the strain was a rod type Gram-negative bacterium having non-motility and did not form spores (see FIG. 1). The cell size was 0.3-0.7 ㎛ in diameter and 0.8-2.5 ㎛ in length. The colony had orange (scarlet) color. The optimum growth temperature of the strain was 25℃, and the strain was never growing under 10℃ or over 40℃ (see FIG. 2). The optimum NaCl concentration for the growth was 1-6%(w/w) and the strain was not growing at all with over 8% (see FIG. 4). The optimum growth pH of the strain was 8 (see FIG. 3).

The strain of the present invention used only D-arabinose and galactose as a carbon source and an energy source for the growth. Neither pentoses, hexoses, sugar alcohols, ogligosaccharides nor other amino acids were used. Starch hydrolysis, cytochrome oxidase and catalase reactions were all positive. But, urease reaction was negative. The

strain could not produce indole from tryptophane, but took advantage of citric acid in deed, confirmed by citric acid test. Denitrification test was also performed. As a result, the strain reduced nitrate to nitrite, but did not reduce nitrite to $N_2$ gas. Besides, the strain did not ferment glucose. The strain of the present invention was also confirmed to be aerobic. DNA G+C composition of the strain was 66.9 mol%. The major non-hydroxyl fatty acid was unsaturated C18:1, and the major hydroxyl fatty acid was C10:0 (3-OH).

In general, a *Paracoccus* genus microorganism is an oxidase and catalase positive, gram-negative bacterium and belongs to α-3-subclass of *Proteobacteria* phylogenetically. Other characteristics of the microorganism were also investigated (see Table 2). As a result, it was confirmed that the strain of the present invention belonged to *Proteobacteria*.

In order to identify the strain of the present invention more concretely, the inventors examined sequence of 16S rDNA. As a result, the nucleotide sequence of 16S rDNA of the strain of

the invention had a high homology with those of *Paracoccus marcusii* and *Paracoccus carotinifaciens*. Nevertheless, the strain of the present invention showed different characteristics, comparing to other *Paracoccus* genus microorganisms including the two above (see Table 4 and Table 5). Therefore, the present inventors confirmed that the strain of the invention was a novel *Paracoccus* genus microorganism and named it '*Paracoccus haeundaensis*'.

The present inventors investigated if carotenoid was generated in the *Paracoccus haeundaensis*, resulting in the confirmation that the strain of the present invention produced β - carotene and astaxanthine (see FIG. 7). The present inventors deposited the *Paracoccus haeundaensis* of the present invention at KCCM (Korean Culture Center of Microorganisms) on January 24, 2003 (Accession No: KCCM - 10460).

*Paracoccus haeundaensis* of the present invention can be effectively used for the production of carotenoid, especially astaxanthine. A method to produce carotenoid using a microorganism of the present invention includes

the steps of culturing *Paracoccus haeundaensis* in a proper medium and collecting astaxanthine from the culture solution. Particularly, the strain of the present invention was cultured in PPES-II medium at 25℃ for 4 days, which was the primary culture, then cells were collected from the culture solution. Organic solvent was added to the cells, which were cultured at 4℃ for overnight, resulting in the elution of astaxanthine. In addition to the above PPES-II medium, LB3 medium (LB medium complemented with 3% NaCl) can be used as a medium for the culture of the strain. It is also preferable to add galactose to a medium in order to induce satisfactory growth of the strain and mass-produce astaxanthine. Methanol, acetone or ethyl ether can be used as an organic solvent used for the purification of astaxanthine from the cultured cells, and methanol is more preferably used. The collection of astaxanthine from the cultured cells can be performed using HPLC (high performance liquid chromatography) or TLC (thin-layer chromatography) by the conventional method known to the people in this field.

The present invention also provides a protein required for biosynthesis of carotenoid and a gene having nucleotide sequences selected from a group consisting of sequences each represented by SEQ. ID. No 5, No 7, No 9, No 11, No 13 and No 15.

The gene above means a gene producing astaxanthine purified from *Paracoccus haeundaensis* (Accession No: KCCM-10460).

It is preferred for a gene coding a protein required for biosynthesis of carotenoid of the present invention to have a sequence selected from a group consisting of sequences each represented by SEQ. ID. No 5, No 7, No 9, No 11, No 13 and No 15. Each protein coded by each gene above has amino acid sequence represented by SEQ. ID. No 6, No 8, No 10, No 12, No 14 and No 16, respectively. 6 genes of the present invention and proteins coded by the same are shown in the below Table 1.

<Table 1>

| Gene | Gene name | Protein | Amino acid sequence |
|---|---|---|---|
| SEQ. ID. No 5 | *crtW* | β -carotene ketolase | SEQ. ID. No 6 |
| SEQ. ID. No 7 | *crtZ* | β -carotene hydroxylase | SEQ. ID. No 8 |

| | | | |
|---|---|---|---|
| SEQ. ID. No 9 | *crtY* | Licopene cyclase | SEQ. ID. No 10 |
| SEQ. ID. No 11 | *crtI* | Phytoene desaturase | SEQ. ID. No 12 |
| SEQ. ID. No 13 | *crtB* | Phytoene synthase | SEQ. ID. No 14 |
| SEQ. ID. No 15 | *crtE* | Geranylgeranyl pyrophosphate synthase | SEQ. ID. No 16 |

Genes provided by the present invention can be effectively used for the production of carotenoid by being inserted in various host cells. Those genes can be used either singly or together (more than 2, at least). For example, a gene coding licopene cyclase and represented by SEQ. ID. No 9 can be used for the production of β -carotene by being inserted in a microorganism containing *crtE*, *crtB* and *crtI* only. And, genes represented by SEQ. ID. No 5 and No 7, coding β -carotene ketolase and β -carotene hydroxylase respectively, can be used for the production of astaxanthine by being inserted in a microorganism producing β -carotene (ex: *Phaffia rhodozyma* ATCC96815).

The present invention further provides a carotenoid synthesis gene containing all the above genes.

A carotenoid synthesis gene is preferred to

have nucleotide sequences represented by SEQ. ID. No 4. The carotenoid synthesis gene (referred as '*crt* gene' hereinafter) of the present invention includes all the carotenoid synthesis genes involved in astaxanthine production process. An organization of *crt* gene of the present invention is presented in FIG. 15. As shown in FIG. 15, the size of *crt* gene of the present invention is 6,223 bp and includes *crtW*, *crtZ*, *crtY*, *crtI*, and *crtB* in that order in 5'→ 3' direction. Each of *Kpn*I, *Sma*I, *Xma*I, *Cla*I, *Hin*dIII and *Bam*HI recognition sequence is located therein. Stop codon of each *crtW*, *crtZ*, *crtY*, *crtI* and *crtB* is overlapped to start codon of the next gene. In particular, *crtE* gene is found as a complementary strand.

The present invention also provides a recombinant vector containing the above carotenoid biosynthesis gene.

A recombinant vector of the present invention was constructed by inserting *crt* gene into a basic vector. Any basic vector that was generally used for gene cloning or expression could be used for the present invention without limitation. And a choice of a vector depended on a host cell. For

example, if *E. coli* is used as a host cell, an *E. coli* specific vector having replication origin of the *E. coli* is preferred. Likewise, if yeast is used as a host cell, a yeast specific vector having replication origin of yeast is preferred. A shuttle vector that has both replication origin of *E. coli* and replication origin of yeast at the same time is also available. In the preferred embodiment of the present invention, the present inventors constructed a recombinant vector containing *crt* gene by using pCR-XL-TOPO vector, which was named 'pCR-XL-TOPO crtfull'.

The present invention also provides a strain prepared by transformed host cells with a recombinant vector containing the above carotenoid biosynthesis gene.

*E. coli* or yeast can be used as host cells of the present invention, and in particular, *E. coli* is preferably selected from a group consisting of XLI-Blue, TOPO, BL21(DE3) codon plus, DH1 and DH5α , but the choice is not always limited thereto. In a preferred embodiment of the present invention, a strain was prepared by transformed BL21(DE3) codon plus, a kind of *E. coli*, with a

recombinant vector 'pCR-XL-TOPO crtfull' which contained *crt* gene represented by SEQ. ID. No 4.

The present invention further provides a method for the production of carotenoid using the carotenoid biosynthesis gene.

The carotenoid producing method of the present invention is comprised of the following steps:

1) Cloning *crt* gene represented by SEQ. ID. No 4;

2) Constructing a recombinant vector in which the gene of the above step 1) was inserted;

3) Transforming a host cell with the recombinant vector of the step 2); and

4) Recovering carotenoids from the culture solution in which a strain transformed with the above recombinant vector was being cultured.

*E. coli* can be used as a host cell. At this time, any *E. coli* strain generally used for the transformation can be used without limitation, but it is preferred to select *E. coli* from a group consisting of XLI-Blue, TOPO, BL21(DE3) codon plus, DH1 and DH5α . In a preferred embodiment of the present invention, BL21(DE3) codon plus was

selected. The choice of a host cell for the invention is not limited to *E. coli*, and yeast is also available.

Particularly, a strain constructed in the above step 1) ~ step 3) was cultured in a growth medium (primary culture). Cells were recovered from the culture solution. Organic solvent was added to the cells, which was further cultured at 4℃ for overnight (secondary culture). At that time, it was possible to add IPTG (isopropyl-beta-D-thiogalactopyranside), an inducer inducing the production of carotenoid, into the culture solution. Carotenoid substrates such as FPP (farnesyl pyrophosphate), GGPP (geranylgeranyl diphosphate) or GPP (geranylpyrophosphate) could also be added. Methanol, acetone or ethyl ether could be used as an organic solvent for the extraction of carotenoid from the culture cells, and methanol was more preferred. Carotenoid was recovered from the culture cells by using HPLC (high performance liquid chromatography) or TLC (thin-layer chromatography) by following the conventional method commonly known for the people in this field.

In a preferred embodiment of the present invention, the inventors measured the amount of astaxanthine produced from a transgenic strain containing *crt* gene. As a result, the produced astaxanthine was 110 μg/g (dry weight), which was far more than that produced by an astaxanthine producing strain '*Paracoccus haeundaensis*' (25 μg/g (dry weight)). Therefore, the method for producing astaxanthine of the present invention makes possible even for a strain which cannot produce astaxanthine itself to mass-produce astaxanthine by using carotenoid biosynthesis gene, so that it facilitates the production of medical supplies and edible pigments as a food additive containing astaxanthine.

In an example of the present invention, a genomic DNA library of *Paracoccus haeundaensis* was constructed in order to clone a gene coding a protein required for carotenoid biosynthesis. The construction of a genomic DNA library was performed by the conventional method commonly known to the people in this field. In particular, a genomic DNA library was constructed by using a cosmid vector in this invention.

In another example of the present invention, 'color complementation' was used for cloning genes, involved in carotenoid biosynthesis, from a genomic DNA library. A microorganism not producing carotenoid (ex: *E. coli*) is given power to generate carotenoid by being transformed with a carotenoid biosynthesis gene that was cloned from a carotenoid producing microorganism (ex: *Paracoccus haeundaensis* of the present invention). For example, *E.coli* could produce β -carotene and cells turned into yellow after being transformed with *crtE, crtB, crtI* and *crtY*. And *E. coli* transformed with *crtE, crtB, crtI, crtY, crtW* and *crtZ* could produce astaxanthine and cells turned into orange. The present inventors cultured a genomic DNA library of *Paracoccus haeundaensis* in a medium supplemented with FPP (farnesyl pyrophosphate), a common substrate of carotenoid. As a result, 13 colonies having orange color were selected. Then, a cosmid vector was isolated from each colony. Nucleotide sequence of DNA insert included in the vector was identified. As a result, it was confirmed that the size of the smallest DNA insert was 6,223 bp. The nucleotide sequence of the DNA insert was represented by SEQ.

ID. No 4.

In another example of the present invention, nucleotide sequences of 6,223 bp long DNA insert assumed to contain a gene producing carotenoid was investigated, resulting in the obtainment of 6 ORFs. Those were analyzed nucleotided on NCBI GenBank. As a result, amino acid sequences translated from each ORF had a high homology with amino acid sequences of 6 enzymes involved in the reaction inducing astaxanthine production from FPP (see FIG. 9 - FIG. 14). From the result, the present inventors confirmed that the DNA insert isolated in this invention had *crt* gene coding a protein necessary for carotenoid biosynthesis (see FIG. 15).

In another example of the present invention, the inventors investigated if *E. coli* not producing carotenoid could produce it by the insertion of *crt* gene coding a relevant protein. First, recombinant vector 'pCR-XL-TOPO-crtfull', in which *crt* gene was inserted, was constructed (see FIG. 16). Then, the prepared recombinant vector was inserted in *E. coli*. Lastly, *E. coli* transformant having orange color was selected. In order to confirm if the transformant could produce

astaxanthine, the transformant was cultured and then cells were collected. Methanol was added to the obtained cells, which were cultured at 4℃ for overnight. Then, supernatant was obtained and optical density was measured at 190-900 nm. As a result, as shown in FIG. 17, original peaks of β-carotene and astaxanthine were confirmed. For more accurate analysis, HPLC assay was performed with some of the supernatant. At that time, β-carotene and astaxanthine purchased from Sigma were used as standard substances. As a result, it was confirmed that a transformant, in which *crt* gene isolated by the present inventors was inserted, could produce β-carotene and astaxanthine. The amount of astaxanthine produced from the transformant of the present invention was 110 μg/g (dry weight). The result suggested that *E. coli* in which *crt* gene was inserted could produce astaxanthine far more than *Paracoccus haeundaensis*, an astaxanthine producing strain, could do (25 μg/g (dry weight)).

## BRIEF DESCRIPTION OF THE DRAWINGS

The application of the preferred embodiments

of the present invention is best understood with reference to the accompanying drawings, wherein:

FIG. 1 is a set of photographs showing the result of observation with a transmission electron microscope on the strain of the present invention in exponential stage, and each bar is 200 nm in length.

FIG. 2 is a graph showing the growth curve of the strain of the present invention depending on temperature.

FIG. 3 is a graph showing the growth curve of the strain of the present invention depending on pH.

FIG. 4 is a graph showing the growth curve of the strain of the present invention depending on culture time.

FIG. 5 is a graph showing the growth curve of the strain of the present invention depending on NaCl concentration.

FIG. 6 is a schematic diagram showing the phylogenetic dendrogram of the *Paracoccus* genus microorganisms of the present invention, based on 16S rDNA sequence data, and a bar represents maximum-parsimony distance (1 nucleotide substitutions per 100 nucleotides).

FIG. 7 is a set of graphs showing the results of scanning on the ranges of optical density changes at 190 ~ 890 nm (A) and at 400 ~ 550 nm (B) of a methanol extract extracted from the culture solution of the strain of the present invention. The peak at 450 nm is the peculiar peak of β -carotene, and the peak at 470 nm is the original peak of astaxanthine.

FIG. 8 is a schematic diagram showing the pathway of carotenoid biosynthesis.

FIG. 9 is a schematic diagram showing the result of comparison between an amino acid sequence of the first ORF (open reading frame) included in a DNA insert isolated from *Paracoccus haeundaensis* and amino acid sequences of β'-carotene ketolase (crtW) isolated from an

*Alcaligenes* genus microorganism and a *Bradyrhizobium* genus microorganism.

P. haeundaensis: *Paracoccus haeundaensis*

Alcaligenes_sp: *Alcaligenes* genus microorganism

Bradyrhizobium_sp: *Bradyrhizobium* genus microorganism

Consensus: corresponding amino acid sequence

FIG. 10 is a schematic diagram showing the result of comparison between amino acid sequences of the second ORF included in a DNA insert isolated from *Paracoccus haeundaensis* and amino acid sequences of β-carotene hydroxylase (crtZ) isolated from an *Alcaligenes* genus microorganism.

P. haeundaensis: *Paracoccus haeundaensis*

Alcaligenes_sp: *Alcaligenes* genus microorganism

Consensus: corresponding amino acid sequence

FIG. 11 is a schematic diagram showing the result of comparison between amino acid sequences of the third ORF included in a DNA insert isolated from *Paracoccus haeundaensis* and amino acid sequences of licopene cyclase (*crtY*) isolated from

an *Flavobacterium* genus microorganism.

P. haeundaensis: *Paracoccus haeundaensis*

Flavobacterium_sp: *Flavobacterium* genus microorganism

Consensus: corresponding amino acid sequence

FIG. 12 is a schematic diagram showing the result of comparison between amino acid sequences of the fourth ORF included in a DNA insert isolated from *Paracoccus haeundaensis* and amino acid sequences of phytoene desaturase (*crtI*) isolated from an *Flavobacterium* genus microorganism.

P. haeundaensis: *Paracoccus haeundaensis*

Flavobacterium_sp: *Flavobacterium* genus microorganism

Consensus: corresponding amino acid sequence

FIG. 13 is a schematic diagram showing the result of comparison between amino acid sequences of the fifth ORF included in a DNA insert isolated from *Paracoccus haeundaensis* and amino acid sequences of phytoene synthase (*crtB*) isolated from an *Flavobacterium* genus microorganism.

P. haeundaensis: *Paracoccus haeundaensis*

Flavobacterium_sp: *Flavobacterium* genus microorganism

Consensus: corresponding amino acid sequence

FIG. 14 is a schematic diagram showing the result of comparison between amino acid sequences of the sixth ORF included in a DNA insert isolated from *Paracoccus haeundaensis* and amino acid sequences of geranylgeranyl pyrophosphate synthase (*crtE*) isolated from an *Flavobacterium* genus microorganism.

P. haeundaensis: *Paracoccus haeundaensis*

Flavobacterium_sp: *Flavobacterium* genus microorganism

FIG. 15 is a schematic diagram showing the organization of *crt* gene isolated from *Paracoccus haeundaensis*.

FIG. 16 is a schematic diagram showing the cleavage map of pCR-XL-TOPO-crtfull vector of the present invention.

FIG. 17 is a set of graphs showing the results of scanning on the ranges of optical

density changes at 190 ~ 890 nm (A) and at 350 ~ 550 nm (B) of a methanol extract extracted from the culture cells of the transformant of the present invention in which *crt* gene was inserted. The peak at 450 nm is the peculiar peak of β - carotene, and the peak at 470 nm is the original peak of astaxanthine.

## EXAMPLES

Practical and presently preferred embodiments of the present invention are illustrative as shown in the following Examples.

However, it will be appreciated that those skilled in the art, on consideration of this disclosure, may make modifications and improvements within the spirit and scope of the present invention.

### Example 1: Sample collection, microorganism separation, cultivation and maintenance of the same

Seawater sample taken from Haeundae shore in

Busan, Korea, was diluted by 1/1000, which was then smeared on a nutrient agar medium (Difco) and cultured at 25℃ for 3 days. Among microorganisms cultured thereby, those having orange or red color were isolated. In order to isolate a microorganism producing carotenoid among strains isolated, all the candidates were cultured in PPES-Ⅱ medium (tripton 1 g/ℓ , bacto-soyton 1 g/ℓ , ferric citrate 0.01 g/ℓ , polypeptone 2 g/ℓ and sodium chloride 3 g/ℓ ). Cells were collected from the culture solution. Carotenoid was extracted by using methanol to identify a final strain producing carotenoid. The isolated microorganism was named 'BC74171$^T$ strain'.

Example 2: Investigation of phenotypic characteristics

In order to investigate morphological and physiological characteristics of the BC74171$^T$ strain of the present invention, experiments were performed as follows.

<2-1> Morphological characteristics

BC74171$^T$ strain of the present invention was cultured in PPES-II medium at 25℃ for 3 days. The cultured cells were suspended in 0.1 M phosphate buffer (pH 7.2). The cells were fixed by 2% glutaraldehyde, washed with 0.05 M cacodylate buffer solution, and fixed again by 1% osmium tetroxide. The fixed cells were dehydrated in ethanol that was later replaced with propylene. Those cells were put into EPON resin and sections were prepared using an ultramicrotome. The prepared sections were observed under a JEM 1200EX-II transmission electron microscope (TEM).

As a result, as shown in FIG. 1, BC74171$^T$ strain of the present invention had a shape of rod and was 0.3-0.7 ㎛ in diameter and 0.8-2.5 ㎛ in length. The strain did not form spores. The mobility of the strain was also observed under an optical microscope by hanging-drop technique (Skerman, V. B. D., *A Guide to the Identification of the Genera of Bacteria, 2nd den. Batimore,* 1967). As a result, BC74171$^T$ strain of the present invention was non-motile. In addition, BC74171$^T$ strain was Gram-negative, and its colonies were flat and had a light orange color.

<2-2> Physiological characteristics

1. Range of temperature

BC74171$^T$ of the present invention was cultured in a nutrition agar medium (Difco) at different temperatures (4, 10, 20, 25, 28, 30, 37, 40 and 50℃) for 10 days to investigate the range of growth temperature. As a result, BC74171$^T$ strain was growing at temperature ranging 20-37℃, and the optimum growth temperature was 25℃ (Table 2 and FIG. 2).

<Table 2>

Characteristics of *Paracoccus haeundaensis*

| Morphological characteristics | | Utilization | |
|---|---|---|---|
| Cell diameter(㎛) | 0.3-0.7 | D-Glucose | - |
| Cell length(㎛) | 0.8-2.5 | Maltose | - |
| Mobility | *- | D-Galactose | *+ |
| Spore formation | - | Sucrose | - |
| Optimum growth temperature(℃) | 25 | Mannitol | - |
| Optimum growth pH | 8 | Cellobiose | - |
| NaCl resistance(%) | 7 | Trehalose | - |

| Product | | Xylose | – |
|---|---|---|---|
| Indole | – | Dulcito | – |
| MR(Methyl red test) | – | Salicin | – |
| VP(Voges-Proskauer test) | – | Adonitol | – |
| Hydrogen sulfide | - | Inositol | – |
| Citrate | + | Arabinose | + |
| Enzyme activity | | Raffinose | – |
| Catalase | + | Rhamnose | – |
| Urease | – | D-Fructose | – |
| Oxidase | + | D-Mannose | – |
| Starch hydrolysis | + | Dimethylformamide | – |
| Denitrification | | Glycerol | – |
| Nitrate → Nitrite | + | L-Glutamic acid | – |
| Nitrite → $N_2$ gas | – | Sorbitol | – |
| Color | +(light orange) | Lactose | – |
| Chemotaxonomic characteristics | | L-Asparagine | – |
| G + C content (mol%) | 66.9 | Acetone | – |
| Non-hydroxyl acids | $C_{18:1}$ | Major carotenoid | Astaxanthine |
| 3-Hydroxyl fatty acids | 3-OH $C_{10:0}$ | | |

*-: Positive response, *+: Negative response

2. Range of pH

BC74171$^T$ of the present invention was cultured in PPES-Ⅱ media each having different pH (3.0, 4.0, 5.0, 6.0, 7.0, 8.0, 9.0, 10.0. 10.5 and 11.0) for 10 days to investigate growth pH range. As a result, the optimum growth pH was 8. The growth of the strain was inhibited or retarded under pH 6 and over pH 10.5 (Table 2 and FIG. 3).

3. Growth

BC74171$^T$ strain of the present invention was shaking-cultured in a PPES-II medium at 25℃ for 10 days to investigate its growth by measuring turbidity. As a result, the growth of the strain was rapidly increased from 30 hours after the culture and slowly decreased from 50 hours after the culture (FIG. 4).

4. Resistance against NaCl

BC74171$^T$ of the present invention was cultured in a trypticase soy broth medium containing 1-10%(w/v) NaCl for 10 days at various temperatures (4, 10, 20, 25, 28, 30, 37, 40 and 50℃). As a result, the optimum NaCl concentration for growth was 1-6%. The growth was retarded with 7% NaCl and was stopped over 8%

(Table 2 and FIG. 5).

5. Utilizing ability of carbon source

In order to examine a capacity to use a carbon source, a micro-plate including 24 substrates listed in the below Table 3 was used. The final concentration of each substrate was adjusted to 1%, which was loaded in a puple broth base (Difro), after being filtered or sterilized by moist heat. The medium was inoculated with the strain of the present invention, then was cultured at 25℃ for 10 days. Yellow meant a positive reaction and purple, which was an original color of the medium, meant a negative reaction.

<Table 3>

Substrates used in the carbon source utilizing ability test and the utilizing ability

| | Substrate | Utilizing ability | | Substrate | Utilizing ability |
|---|---|---|---|---|---|
| 1 | D-Glucose | *- | 13 | Arabinose | + |
| 2 | Maltose | - | 14 | Raffinose | - |
| 3 | D-Galactose | *+ | 15 | Rhamnose | - |
| 4 | Sucrose | - | 16 | D-Fructose | - |
| 5 | Mannitol | - | 17 | D-Mannose | - |

| | | | | | |
|---|---|---|---|---|---|
| 6 | Cellobiose | - | 18 | Dimethylformamide | - |
| 7 | Trehalose | - | 19 | Glycerol | - |
| 8 | Xylose | - | 20 | L-Glutamic acid | - |
| 9 | Dulcitol | - | 21 | Sorbitol | - |
| 10 | Salicin | - | 22 | Lactose | - |
| 11 | Adonitol | - | 23 | L-Asparagine | - |
| 12 | Inositol | - | 24 | Acetone | - |

*-: Not used (positive reaction),

+: Used (negative reaction)

As a result, as shown in Table 3, BC74171$^T$ strain of the present invention used only D-arabinose and galactose as a carbon source, and no other substrates were used.

6. Activity of starch hydrolysis

In order to investigate if the strain of the present invention could hydrolyze starch, a starch agar was inoculated with the strain, then, cultured for 10 days. An activity of starch hydrolysis was measured by the method of Cowan, S. T. & Steel, K. J. (Cowan, S. T. & Steel, K. J., *Manual for the Identification of Medical Bacteria. London: Cambridge University Press*, 1965). As a

result, the strain of the present invention was confirmed to have an activity of starch hydrolysis (Table 2).

7. Indole test (tryptophanase activity)

In order to investigate if the strain of the present invention could produce indole by decomposing tryptophane, the present inventors took advantage of the method of Cappuccino *et al.* (Cappuccino J. G. and Sherman, N. In *Microbiology: a laboratory manual (6th)* 2001), resulting in the confirmation that indole was not produced from tryptophane (Table 2).

8. Acid production (fermentation) from carbohydrates

Acid production was investigated by the method of Hughm *et al.* (Hughm *et al.*, *J. Bacteriol.*, 66: 24-26, 1953). Methyl Red test and Voges-Proskauer test were performed and the results were all negative (Table 2).

9. Citric acid test

In order to investigate if the strain of the present invention could be growing by using citric

acid as a carbon source, the method of Cappuccino *et al.* (Cappuccino J. G. and Sherman, N. In *Microbiology: a laboratory manual (6th)* 2001) was performed and the result was positive (Table 2).

10. Catalase activity

The catalase activity was investigated by observing bubbles in 3% hydtogen peroxide solution by following the method of Harker *et al.* (Harker *et al., J. Clin. Microbiol.*, 2: 463-464, 1975). As a result, the catalse activity of the strain of the present invention was confirmed to be positive (Table 2).

11. Oxidase activity

The oxidase activity was investigated by using 1% p-aminodimethylaniline oxalate as a substrate by following the method of Cappuccino *et al.* (Cappuccino J. G. and Sherman, N. In *Microbiology: a laboratory manual (6th)* 2001). As a result, the cytochrome oxidase activity of the strain of the present invention was confirmed to be positive (Table 2).

12. Urease activity

The urease activity was measured by the method of Lanyi (Lanyi, B. *Methods Microbiol.*, 19: 1-67, 1987). As a result, the urease activity of the strain of the present invention was confirmed to be negative (Table 2).

13. Denitrification

In order to observe denitrification, gas production and growth were investigated through stab culture using a nutritive medium containing 0.1% (w/v) agar, by following the method of Cappuccino *et al.* (Cappuccino J. G. and Sherman, N. In *Microbiology: a laboratory manual (6th)* 2001). As a result, BC74171$^T$ strain of the present invention reduced nitrate to nitrite. But, the strain could not reduce nitrite to $N_2$ gas (Table 2).

14. Hydrogen sulfide production test

In order to investigate if the strain of the present invention could produce hydrogen sulfide from amino acid containing sulfur like cysteine or a substrate like inorganic sulfur compounds, hydrogen sulfide production test was performed using TSI (Triple Sugar Iron Agar) medium

(Cappuccino J. G. and Sherman, N. In *Microbiology: a laboratory manual (6th)* 2001). The result was negative (Table 2).

Example 3: Fatty acid analysis

BC74171$^T$ strain of the present invention was cultured on trypticase soy agar medium (pH 8.0) supplemented with 2% NaCl at 25℃ for 2 days. Then, harvesting, saponification, methylation and extraction of FAMEs (fatty acid methyl esters) were performed with the cells cultured by the method of Sasser (Sasser, M. *In Methods in Phytobacteriology*, 199-204, 1990). GC-MS was performed following the method of Lipske *et al.* (Lipski *et al.*, *Syst. Appl. Microbiol.*, 20: 448-457, 1997), leading to the investigation of FAMEs. As a result, according to the confirmed fatty acid profile ($C_{18:1}$, 84.32 %; $C_{18:0}$, 7.79 %; $C_{10:0}$ (3-OH), 2.06 %; $C_{12}$:1cis5 2.0 %; $C_{14:0}$ (3-OH), 1.47 %; $C_{17:0}$, 0.80 %; $C_{16:0}$, 0.78 %; and unknown peak, 0.78 %), the strain of the present invention was confirmed to belong to α -subclass of *Proteobacteria*. And a major hydroxyl fatty acid was $C_{10:0}$ (3-OH) (Table 2).

Example 4: Base composition of DNA

The base composition of genomic DNA of BC74171$^T$ strain of the present invention was identified by the method of Tamaoka *et al.* (Tamaoka, J. & Komagata, K. *FEMS Microbiol. Lett.*, 25: 125-128, 1984). Particularly, the genomic DNA was extracted from BC74171$^T$ strain of the present invention by following the conventional method known to the people in this field. The extracted genomic DNA was hydrolyzed. The obtained nucleotides were analyzed by HPLC (reverse-phase HPLC). As a result, G+C content of DNA of BC74171$^T$ strain was 66.9 mol% (Table 2).

Example 5: 16S rDNA sequencing and phylogenetic dendrogram analysis

Genomic DNA was extracted from BC74171$^T$ strain of the present invention by the method of Rainey *et al.* (Rainey *et al.*, *Syst. Appl. Microbiol.*, 15: 197-202, 1992). 16S rDNA was PCR amplified by using primers represented by SEQ. ID. No 1 and no 2. PCR was performed as follows. 20

pmol each of two primers, 10 ng of genomic DNA, 1 unit of Taq polymerase and 10x buffer solution□(with $MgCl_2$) were mixed to make the reacting solution. The prepared reacting solution was pre-denatured at 94℃ for 5 minutes, followed by 25 cycles of denaturation at 94℃ for 1 minute, annealing at 56℃ for 30 seconds, polymerization at 72℃ for 90 seconds, and final extension at 72℃ for 10 minutes. The PCR product was cloned into pGEM-T vector (Promega) and isolated DNA clones were reacted with ABI PRISM™ staining reagent (Perkin Elmer, USA). Nucleotide sequence was determined by using ABI 377 genetic analyzer (Perkin Elmer, USA).

As a result, 16S rDNA of BC74171$^T$ strain of the present invention was 1451 bp long and had nucleotide sequences represented by SEQ. ID. No 3. The nucleotide sequence of 16S rDNA of BC74171$^T$ strain was analyzed by using BLASTN and BLASTX of NCBI GenBank database. As a result, the nucleotide sequence of 16S rDNA of BC74171$^T$ strain showed 99.8% and 99.6% homology each with those of *Paracoccus marcusii* and *Paracoccus carotinifaciens* which have been known to produce astaxanthine among various *Paracoccus* genus microorganisms.

Thus, BC74171$^T$ strain of the present invention was confirmed to be a *Paracoccus* genus strain. And the characteristics of the strain of the present invention were compared with those of various *Paracoccus* genus microorganisms. The characteristics of other *Paracoccus* genus microorganism listed in the below Table 4 (2-15) were the results of investigations by Harker *et al.* (Harker *et al.*, *Int. J. Syst. Bacteriol.*, 48: 543-548, 1998), Lipski *et al.* (Lipski *et al.*, *Syst. Appl. Microbiol.*, 20: 448-457, 1997), Tsubokura *et al.* (Tsubokura *et al.*, *Int. J. Syst. Bacteriol.*, 49: 277-282, 1999), Kelly *et al.* (Kelly *et al.*, The genus *Paracoccus*. In *The Prokaryotes*, http://www.prokaryotes.com., Edited by M. Dworkin, S. Falkow, E. Rosenberg, K. H. Schleifer & E. Stackebrandt. New York: Springer., 2000) and Doronina *et al.* (Doronina *et al.*, *Int. J. Syst. Evol. Bacteriol.*, 52: 679-6, 2002). As shown in Table 4, BC74171$^T$ strain of the present invention did not share any characteristics with any other *Paracoccus* genus microorganism.

<Table 4>

Comparison of the characteristics of the strain of

the present invention and those of other *Paracoccus* genus microorganisms

| | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Mobility | | − | − | − | − | − | + | − | − | − | − | − | − | − | − | + |
| Growth in 6% NaCl | | + | − | NR | − | − | NR | − | − | − | W | NR | NR | − | NR | NR |
| Usefulness | Glucose | − | + | − | + | + | + | + | − | + | + | + | + | − | + | + |
| | Arabinose | + | + | − | − | − | − | − | − | + | + | + | − | − | − | − |
| | Glycerol | − | + | − | + | + | NR | + | − | + | + | + | + | − | − | + |
| | Sucrose | − | − | − | − | − | NR | + | − | − | + | + | + | − | − | + |
| | Mannitol | − | + | − | − | + | + | + | − | + | + | + | + | − | + | + |
| | Fructose | − | + | − | − | + | NR | + | − | + | + | + | + | − | + | + |
| Denitrification | | − | − | + | − | − | − | + | + | + | − | − | + | + | + | + |
| Urease activity | | − | + | NR | − | − | − | + | − | + | NR | + | − | NR | − | − |
| Yellow color | | + | − | NR | − | − | + | − | − | − | + | − | − | − | − | − |
| G + C content (mol%) | | 66.9 | 64-66 | NR | 63 | 67 | 67 | 64-67 | 71 | 62.5 | 66 | 67 | 66 | 68.5-70 | 66.5-67.6 | 67-68 |

*1: The strain of the present invention

2: *P. alcaliphibus* (JCM 7364T)

3: *P. alkenifer* (DSM 11593T)

4: *P. aminophilus* (JCM 7686T)

5: *P. aminovorans* (JCM 7685T)

6: *P. carotinifaciens* (IFO 16121T)

7: *P. denitrificans* (ATCC 17741T)

8: *P. kocurii* (JCM 7684T)

9: *P. kondratievae* (VKM B-2222T)

10: *P. marcusii* (DSM 11574T)

11: *P. methylutens* (VKM B-2164T)

12: *P. pantotrophus* (ATCC 35512T)

13: *P. solventivorans* (DSM 6637T)

14: *P. thiocyanatus* (IAM 12816T)

15, *P. versutus* (ATCC 25364T)

*+: Positive reaction -: Negative reaction

W: Weak reaction) NR: Not reported

The characteristics of BC74171$^T$ strain of the present invention were different in many ways even with *Paracoccus marcusii* and *Paracoccus carotinifaciens* showing the highest 16S rDNA sequence homology. More precisely, as shown in the below Table 5, *Paracoccus marcusii* (Harker *et al.*, *Int. J. Syst. Bacteriol.*, 48: 543-548, 1998) was a coccus or a short rod type forming a short chain. On the contrary, BC74171$^T$ strain of the

present invention was a rod type not forming a chain. *Paracoccus carotinifaciens* (Tsubokura *et al.*, *Int. J. Syst. Bacteriol.*, 49: 277-282, 1999) was also a rod type not forming a chain but had flagella. However, BC74171$^T$ strain of the present invention had not flagella. The strain of the present invention did not use any of glucose, mannitol, maltose or mannose, but the other two strains did use them. Besides, BC74171$^T$ strain of the present invention decomposed starch. But, the two other strains did not.

<Table 5>

Comparison of the characteristics of the strain of the present invention and those of other *Paracoccus* genus microorganisms

| Characteristics | | BC74171$^T$ | *Paracoccus marcusii* | *Paracoccus carotinifaciens* |
|---|---|---|---|---|
| Cell morphology | | Rod type | Coccus or short rod type | Rod type |
| Mobility | | −* | − | +* |
| Useful nesses | Glucose | − | + | + |
| | Maltose | − | + | + |
| | Mannitol | − | + | + |
| | Arabinose | + | + | − |

| | | | |
|---|---|---|---|
| Citric acid | + | + | − |
| Mannose | − | NR* | + |
| Starch hydrolysis | + | − | − |
| Color | + | + | + |

*+: Positive reaction, −: Negative reaction, NR: Not reported

Considering all the above results, BC74171$^T$ strain of the present invention was confirmed to be a novel microorganism belonging to *Paracoccus* genus. The present inventors named the BC74171$^T$ strain as '*Paracoccus haeundaensis*' and deposited it at KCCM (Korean Culture Center of Microorganisms), on January 24, 2003 (Accession No: KCCM-10460). Systematic position of the *Paracoccus haeundaensis* of the present invention was shown in FIG. 6. Phylogenetic dendrogram was made out by using Treeview program. Bootstrap analysis (1000 replications) was performed using a method measuring distance and parsimony (Agnes Groisillier and Aline Lonvaud-Funel, *International Journal of Systematic Bacteriology*, 49: 1417-1428, 1999).

Example 6: Astaxanthine production in the strain of the present invention

The strain of the present invention was cultured in 50 ㎖ of medium (yeast extract 1%, tryptone 0.5%, NaCl 3%) at 25℃ for 6 days. The cells were collected by centrifugation (13,000 rpm). The collected cells were suspended in 20 ㎖ of methanol, which were then cultured overnight at 4℃. Centrifugation was performed again with 13,000 rpm to obtain supernatant. Optical density was measured at 190-900 nm.

As a result, as shown in FIG. 7, peaks were observed both at 450 nm and at 470 nm, and those peaks were confirmed to be original peaks of β-carotene and astaxanthine.

For the accuracy of analysis, 1 ㎖ of the supernatant was taken and filtered by a 0.45 ㎛ filter, followed by HPLC analysis (column: 4.6× 250 mm, uBondapak C18, Waters, Milford, MA; mobile phase: acetonitrile-methanol-water(49:44:7 v/v), Flow: 10 ㎖/min, Detector: 470 nm). β-carotene and astaxanthine purchased from Sigma were used as a standard substance. As a result, the strain of the present invention did produce astaxanthine and

β -carotene. The amount of astaxanthine produced from the strain was 25 μg/g (dry weight). Therefore, it was confirmed that bright orange color of the strain was caused by carotenoid biosynthesis in a cell and a major pigment accumulated in the cell was astaxanthine.

Example 7: Preparation of genomic DNA for the cloning of a carotenoid biosynthesis gene

*Paracoccus haeundaensis* (KCCM-10460) was cultured in a PPES-II medium (tripton 1 g/ℓ, bacto-soyton 1 g/ℓ, ferric citrate 0.01 g/ℓ, polypepton 2 g/ℓ and NaCl 3 g/ℓ) at 25℃ for 10 days. Then, the culture solution was centrifuged at 13,000 rpm to collect cells. In order to isolate genomic DNA from cells, the cells were suspended in STE buffer solution (10 mM Tris, 1 mM EDTA, 100 mM NaCl, pH 8.0), which was further reacted at 68℃ for 15 minutes. The cells obtained from centrifugation were re-suspended in solution I (50 mM glucose, 25 mM Tris, 10 mM EDTA, pH 8.0). 5 mg/mℓ of lysozyme and 100 μg/mℓ of RNase A were added, followed by reaction at 37℃ for 1 hour. Then, 250 μg/mℓ of proteinase K was added,

followed by reaction at 37℃for 3 hours. N-lauroylsarcosine was added by 1% of total volume, followed by reaction at 37℃. Genomic DNA was purified by phenol-chloroform extraction method. Extraction was performed by adding same volume of phenol-chloroform and genomic DNA was precipitated by adding 100% ethanol twice as much as the whole volume. The precipitated DNA was washed with 70% ethanol. TE buffer solution was added to dissolve the precipitated DNA completely at 65℃, which was then ready to be used.

Example 8: Construction of genomic DNA library

<8-1> Preparation of a cosmid vector

*Xba* I (9 U/㎍) restriction enzyme was added to 25 ㎍ of a cosmid vector (SuperCos 1 Cosmid Vector, Stratagene), making the total volume of reacting solution 200 ㎕, and the solution was reacted at 37℃ for 1 hour, followed by digestion. Vector DNA was purified by phenol-chloroform extraction method, which was precipitated by 100% ethanol. For dephosphorylation of the vector digested with *Xba* I, CIAP enzyme (Promega) was added, followed by reaction at 37℃ for 30 minutes.

Then, the purified vector was reacted with *Bam*H I (5 U/$\mu$g) at 37℃ for 1 hour. Extraction was performed by phenol-chloroform extraction method, and precipitation was induced by using ethanol. The precipitate was dissolved in TE buffer solution, making the concentration 1 $\mu g/\mu \ell$.

<8-2> Construction of genomic DNA library

100 $\mu$g of *Paracoccus haeundaensis* genomic DNA obtained in the above Example 7 was treated with Sau3A I (10 U), inducing a partial enzyme reaction. Upon completing the enzyme reaction, 0.5 M EDTA was added. Then, genomic DNA was separated by phenol-chloroform extraction method, which was precipitated by 100% ethanol. Genomic DNA, a product of a partial enzyme reaction, was dissolved in TE buffer solution, and treated with CIAP enzyme by the same method as used in the above Example 7, leading to dephosphorylation. DNA was separated again by phenol-chloroform extraction method. For the ligation with a cosmid vector prepared in the above Example <8-1>, T4 ligase (Promega) and 10x ligase buffer (Promega) were added to the isolated DNA, followed by reaction at 12℃ for 18 hours. After completing

the reaction, an *E. coli* strain XL1-Blue (Stratagene) was transformed with the ligation mixture, resulting in the construction of a genomic DNA library.

Example 9: Examination and analysis of a transgenic strain containing a color-producing gene

FPP (Sigma), one of common substrates included in carotenoids, was added to a LB agar medium by 1% out of total volume. A genomic library constructed in the above Example 8 was smeared on the plate, which was cultured at 37℃. Among cultured colonies, the ones having orange color were selected (13 out of about 2000 colonies). Cosmid vectors were isolated from those 13 colonies selected above. Then, primer working sequencing was performed to identify nucleotide sequences of DNA fragments inserted in each cosmid vector. Identification of the nucleotide sequences was committed to GenoTech Corp., Korea.

As a result, the smallest DNA insert of all fragments inserted in the cosmid vectors was 6,223 bp long, and had nucleotide sequences represented by SEQ. ID. No 4. A cosmid vector containing the nucleotide sequence represented by SEQ. ID. No 4 was named 'COSCRT'.

Example 10: Sequence analysis of a DNA insert containing a carotenoid biosynthesis related gene

Sequence analysis of a DNA insert obtained in the above Example 9 was performed using NCBI ORF Finder program (http://www.ncbi.nlm.nih.gov/gorf/gorf.html) to analyze ORF.

As a result, 6 ORFs were included in a DNA insert obtained in the above Example 3. Each ORF showed a relevant homology with nucleotide sequences of *crtW* coding β-carotene ketolase, *crtZ* coding β-carotene hydroxylase, *crtY* coding licopene cyclase, *crtI* coding phytoene desaturase, *crtB* coding phytoene synthase and *crtE* coding geranylgeranyl pyrophosphate synthase. All of those enzymes were confirmed to be involved in carotenoid biosynthesis.

The comparison of homology between amino acid sequences translated from each ORF and amino acid sequences of each carotenoid biosynthesis enzyme isolated from *Alcaligenes* sp. (Misawa *et al.*, *Biochem. Biophys. Res. Commun.*, 209(3): 867-876, 1995), *Bradyrhizobium* sp. (Hannibal *et al.*, *J. Bacteriol.*, 182(13): 3850-3853, 2000) and *Flavobacterium* sp. (Pasamotes *et al.*, *Gene*, 185(1): 35-41, 1997) was shown in FIG. 9 and FIG. 14. The nucleotide sequences of 6 ORF genes cloned by the present inventors and amino acid sequences translated therefrom were each represented by SEQ. ID. No 5 ~ No 16. In particular, *crtW*, *crtZ*, *crtY*, *crtI* and *crtB* genes were each represented by SEQ. ID. No 5, No 7, No 9, No 11, No 13 and No 15 and amino acid sequences of each gene above were represented by SEQ. ID. No 6, No 8, No 9, No 12, No 13 and No 16, respectively.

It was confirmed from the above results that *crt* gene involved in biosynthesis of carotenoid was included in a DNA fragment inserted in the cosmid vector above.

The composition of *crt* gene of the present invention was shown in FIG. 15. As shown in FIG. 15, termination codon and start codon were

overlapped in each *crtW*, *crtZ*, *crtY*, and *crtI*. Especially, *crtE* gene seemed to have a directionality of complementary strands, and had each of recognition sequences of *Kpn*I, *Xma*I, *Sma*I, *Cla*I, *Hind*III and *Bam*H in its sequences.

Example 11: Expression of *crt* gene in *E. coli*

In order to investigate if carotenoid could be produced by a protein expressed by *crt* gene of *Paracoccus haeundaensis* isolated in the above Example 9, *crt* gene was first amplified by PCR using HL premix (Bioneer). At that time, oligoneucleotide primers represented by SEQ. ID. No 17 and No 18 were used. All the PCRs were performed as follows; pre-denaturation at 94℃ for 5 minutes, denaturation at 94℃ for 30 seconds, annealing at 66℃ for 30 seconds, polymerization at 72℃ for 6 minute, 25 cycles from denaturation to polymerization, and final extension at 72℃ for 20 minutes. The PCR product was inserted in Topo-XL-vector (Invitrogen), which was used for transduction of *E. coli*. At that time, XL1-Blue (Stratagene), TOPO (Invitrogen), BL21(DE3) codon plus (Stratagene), DH1 (Takara) and DH5α (Takara)

were used as *E. coli* of the present invention. As a result, all transformants of BL21(DE3), XL1-Blue, BL21(DE3) codon plus were confirmed to have orange color. Thereafter, each transformant of *E. coli* was selected for culture. As a result, transformed BL21(DE3) codon plus produced astaxanthine most.

The present inventors inserted *crt* gene having nucleotide sequences represented by SEQ. ID. No 4, isolated in the above Example 9, into pCR-XL-TOPO vector (Invitrogen), a gene expression vector, which was then named 'pCR-XL-TOPO-crtfull' (FIG. 16). BL21(DE3) codon plus cells were transfected with the vector 'pCR-XL-TOPO-crtfull, followed by selection of strains having orange color. The selected strains were cultured in 50 mℓ of LB medium at 37℃ for 8 hours. The culture solution was centrifuged at 13,000 rpm, then supernatant was discarded and cells were collected. 20 mℓ of methanol was added to the collected cells. After vortexing, the cells were cultured at 4℃ for overnight. Centrifugation was performed (13,000 rpm) to obtain supernatant. In order to confirm whether carotenoid was generated, optical

density was measured at 190-900 nm and 400-550 nm. As a result, peaks were observed at 450 nm and 470 nm, which were confirmed to be original peaks of β -carotene and astaxanthine (FIG. 17).

For the accuracy of analysis, 1 mℓ of the above supernatant was obtained and filtered with a 0.45 μm (pore size) filter, followed by HPLC analysis (column: 4.6x250mm, uBondapak C18, Waters, Milford, MA; mobile phase: acetonitrile-methanol-water(49:44:7 v/v), Flow: 10ml/min, Detector: 470nm). At that time, β -carotene and astaxanthine, purchased from Sigma, were used as a standard substance. As a result, the substances produced from the strain of the present invention were confirmed to be astaxanthine and β -carotene.

Astaxanthine produced from the strain of the present invention was also quantified, resulting in the production of 110 μg/g (dry weight). For that result, neither inducer nor carotenoid substrate was added, suggesting that β -carotene and astaxanthine could be produced in *E. coli* only by nucleotide sequences having 6,223 bp isolated in the present invention. The amount of astaxanthine quantified above was far more than that produced (25 μg/g) by *Paracoccus haeundaensis*

(Accession No: KCCM-10460).

## INDUSTRIAL APPLICABILITY

As explained hereinbefore, the present inventors have isolated and identified a novel *Paracoccus* genus microorganism mostly producing astaxanthine among many carotenoids, and also have cloned 6 genes coding proteins involved in carotenoid biosynthesis and *crt* gene containing the same from the above microorganism. The present inventors also have confirmed that carotenoid can be produced even in *E. coli* not producing carotenoid, by using *crt* gene. Therefore, a gene of the present invention and a microorganism producing the gene can be effectively used for the production of carotenoids such as β-carotene and astaxanthine which are available for making food, medicines and beauty stuffs.

Those skilled in the art will appreciate that the conceptions and specific embodiments disclosed in the foregoing description may be readily

utilized as a basis for modifying or designing other embodiments for carrying out the same purposes of the present invention. Those skilled in the art will also appreciate that such equivalent embodiments do not depart from the spirit and scope of the invention as set forth in the appended claims.

**What is claimed is**

1. A gene coding a protein involved in carotenoid biosynthesis, which has nucleotide sequences selected from a group consisting of nucleotide sequences represented by SEQ. ID. No 5, No 7, No 9, No 11, No 13 and No 15.

2. The gene as set forth in claim 1, wherein the gene has nucleotide sequences of *crtW* coding β -carotene ketolase and represented by SEQ. ID. No 5.

3. The gene as set forth in claim 1, wherein the gene has nucleotide sequences of *crtZ* coding β -carotene hydroxylase and represented by SEQ. ID. No 7.

4. The gene as set forth in claim 1, wherein the gene has nucleotide sequences of *crtY* coding licopene cyclase and represented by SEQ. ID. No 9.

5. The gene as set forth in claim 1, wherein the gene has nucleotide sequences of *crtI* coding

phytoene desaturase and represented by SEQ. ID. No 11.

6. The gene as set forth in claim 1, wherein the gene has nucleotide sequences of *crtB* coding phytoene synthase and represented by SEQ. ID. No 13.

7. The gene as set forth in claim 1, wherein the gene has nucleotide sequences of *crtE* coding geranylgeranyl pyrophosphate synthase and represented by SEQ. ID. No 15.

8. A *crt* gene containing all the genes of claim 2 ~ claim 7 and represented by SEQ. ID. No 4.

9. A protein encoded by the gene of claim 1, which has nucleotide sequences selected from a group consisting of nucleotide sequences represented by SEQ. ID. No 6, No 8, No 10, No 12, No 14 and No 16.

10. A recombinant vector containing the *crt* gene of claim 8.

11. The recombinant vector as set forth in claim 10, wherein the vector is pCR-XL-TOPO-crtfull having a cleavage map represented in FIG. 16.

12. An *E. coli* transformant transformed with the recombinant vector of claim 11.

13. A method for producing carotenoid comprising the following steps:

1) Cloning the *crt* gene of claim 8;

2) Constructing a recombinant vector in which the *crt* gene of the above step 1) was inserted;

3) Transfecting a host cell with the recombinant vector of the step 2); and

4) Recovering carotenoids from the culture cells in which a strain transformed with the above recombinant vector was being cultured.

14. The method as set forth in claim 13, wherein the recombinant vector is that of claim 11.

15. The method as set forth in claim 13, wherein the host cell is *E. coli* or yeast.

16. The method as set forth in claim 13, wherein the recovery of carotenoids is performed from the culture cells in which the *E. coli* was being cultured.

17. The method as set forth in claim 13, wherein the cartenoid is β -carotene or astaxanthine.

18. A *Paracoccus haeundaensis* producing astaxanthine, which has a 16S rDNA nucleotide sequence represented by SEQ. ID. No 3.

19. The *Paracoccus haeundaensis* as set forth in claim 18, wherein the strain is represented by accession No: KCCM-10460.

FIGURES

FIG. 1

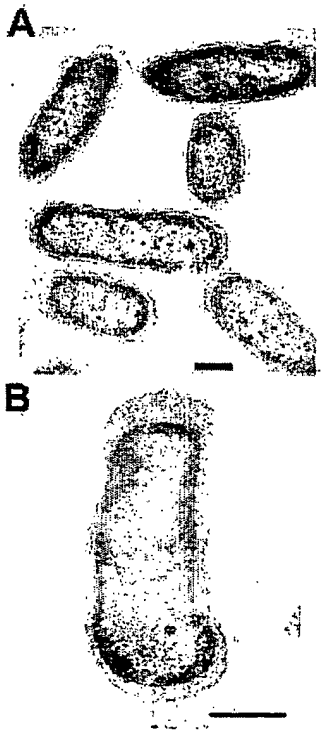


FIG. 2

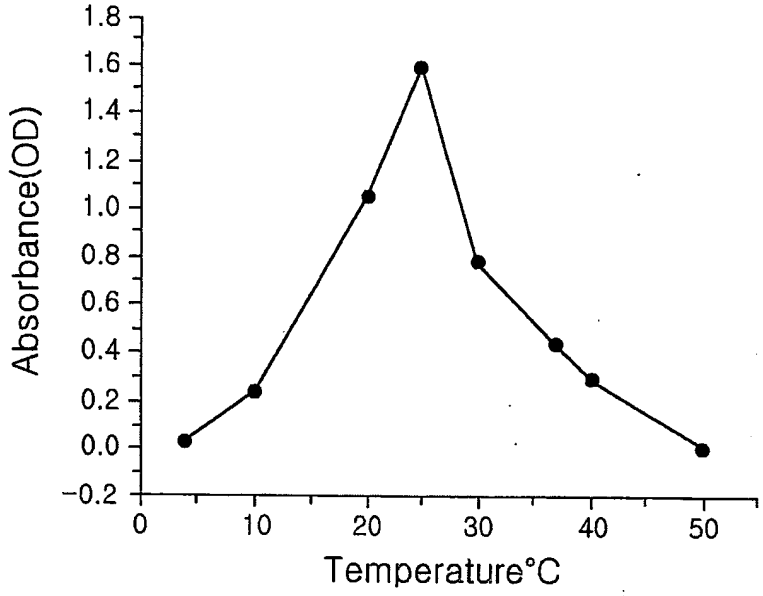

Absorbance(OD)
1.8
1.6
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
-0.2
0
10
20
30
40
50
Temperature°C

FIG. 3

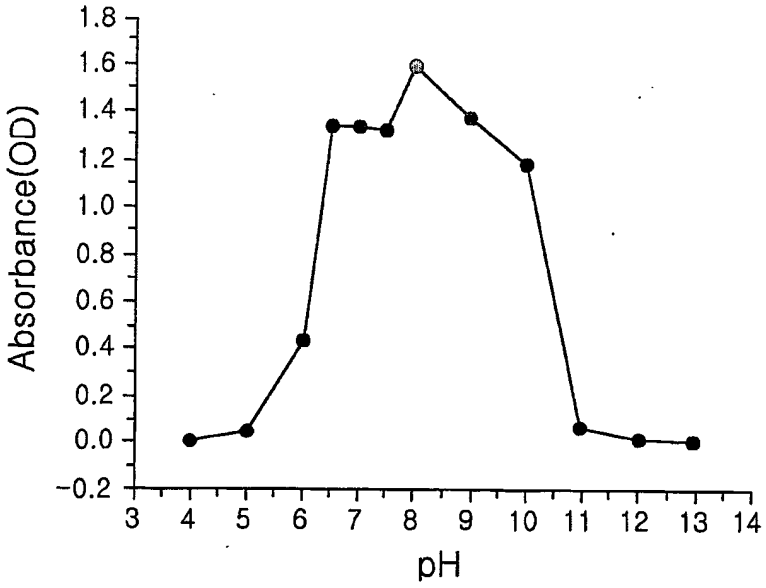

Absorbance(OD)
1.8
1.6
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
-0.2
3
4
5
6
7
8
9
10
11
12
13
14
pH

FIG. 4

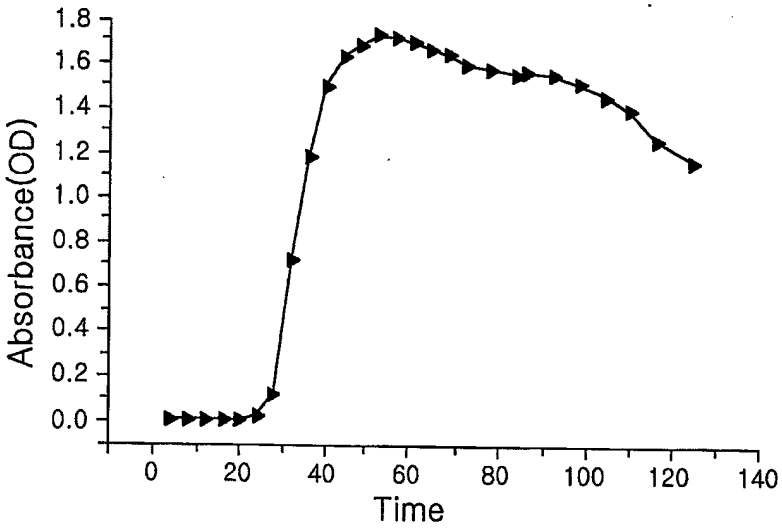
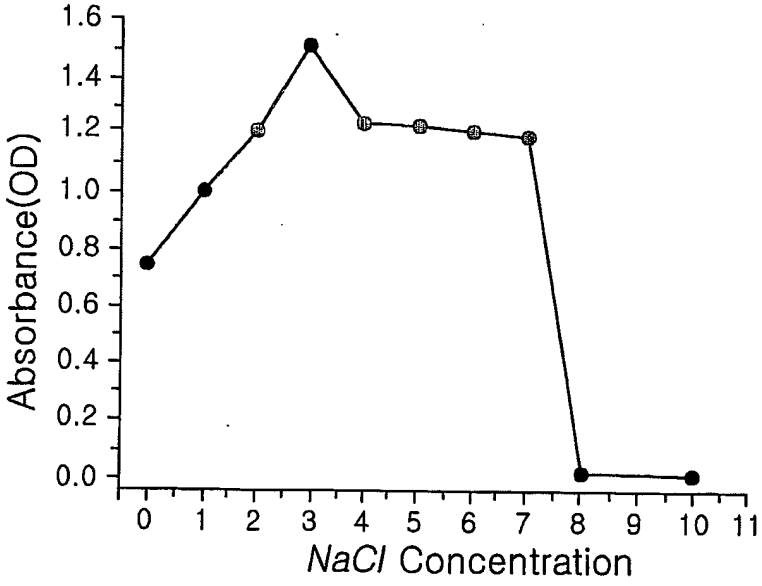

Absorbance(OD)
1.8
1.6
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
0
20
40
60
80
100
120
140
Time

FIG. 5


Absorbance(OD)
1.6
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
0
1
2
3
4
5
6
7
8
9
10
11
NaCl Concentration

4/11

FIG. 6

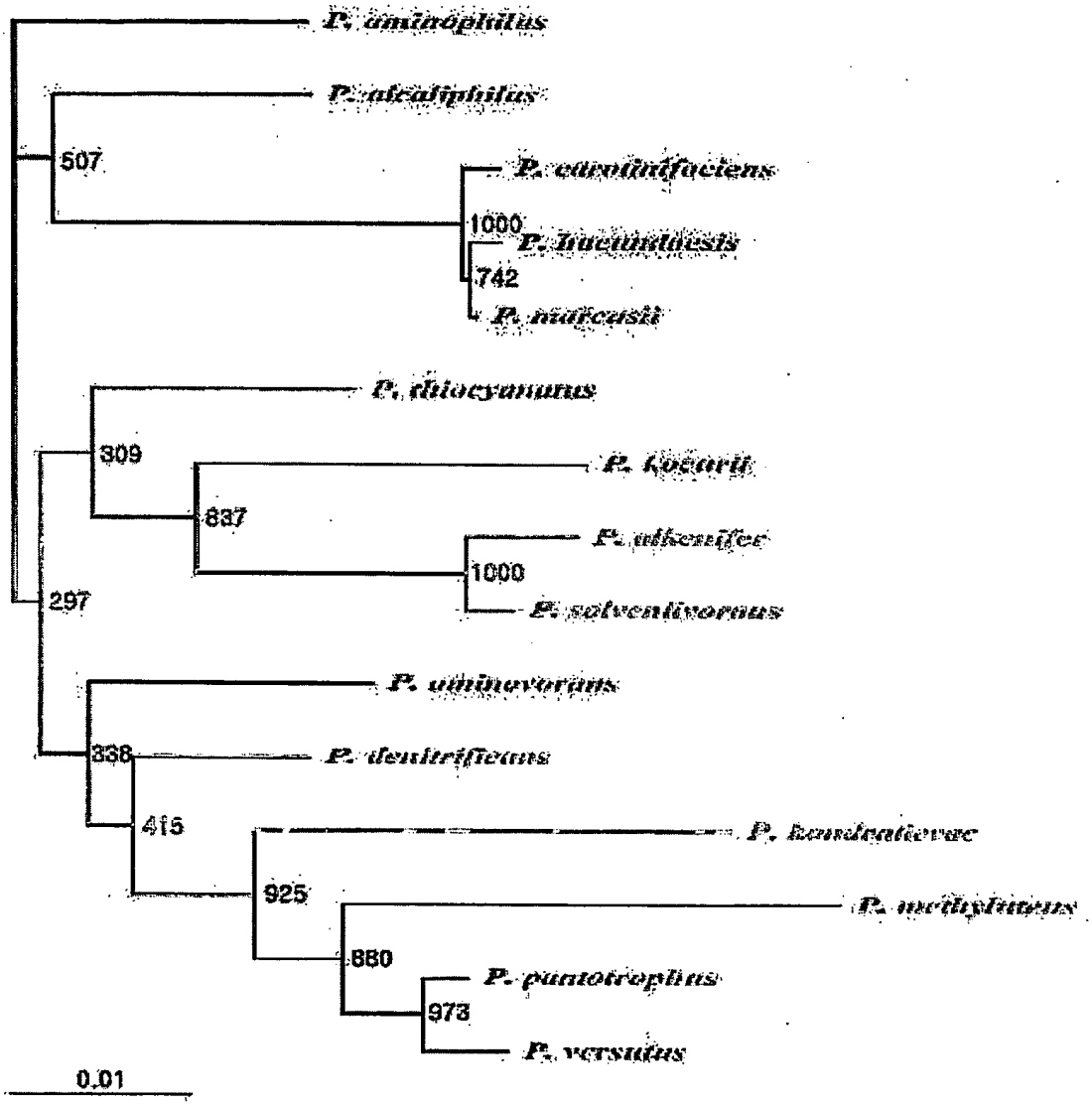

P. aminophilus
P. alcaliphilus
507
P. cumulifaciens
1000
P. haeundaesis
742
P. naejeonsis
P. thiocyanatus
309
P. koreensis
837
P. alkenifer
1000
297
P. solventivorans
P. uninevorans
338
P. denitrificans
415
P. kondratievae
925
P. methylovorus
880
P. pantotrophus
973
P. versutus
0.01

FIG. 7

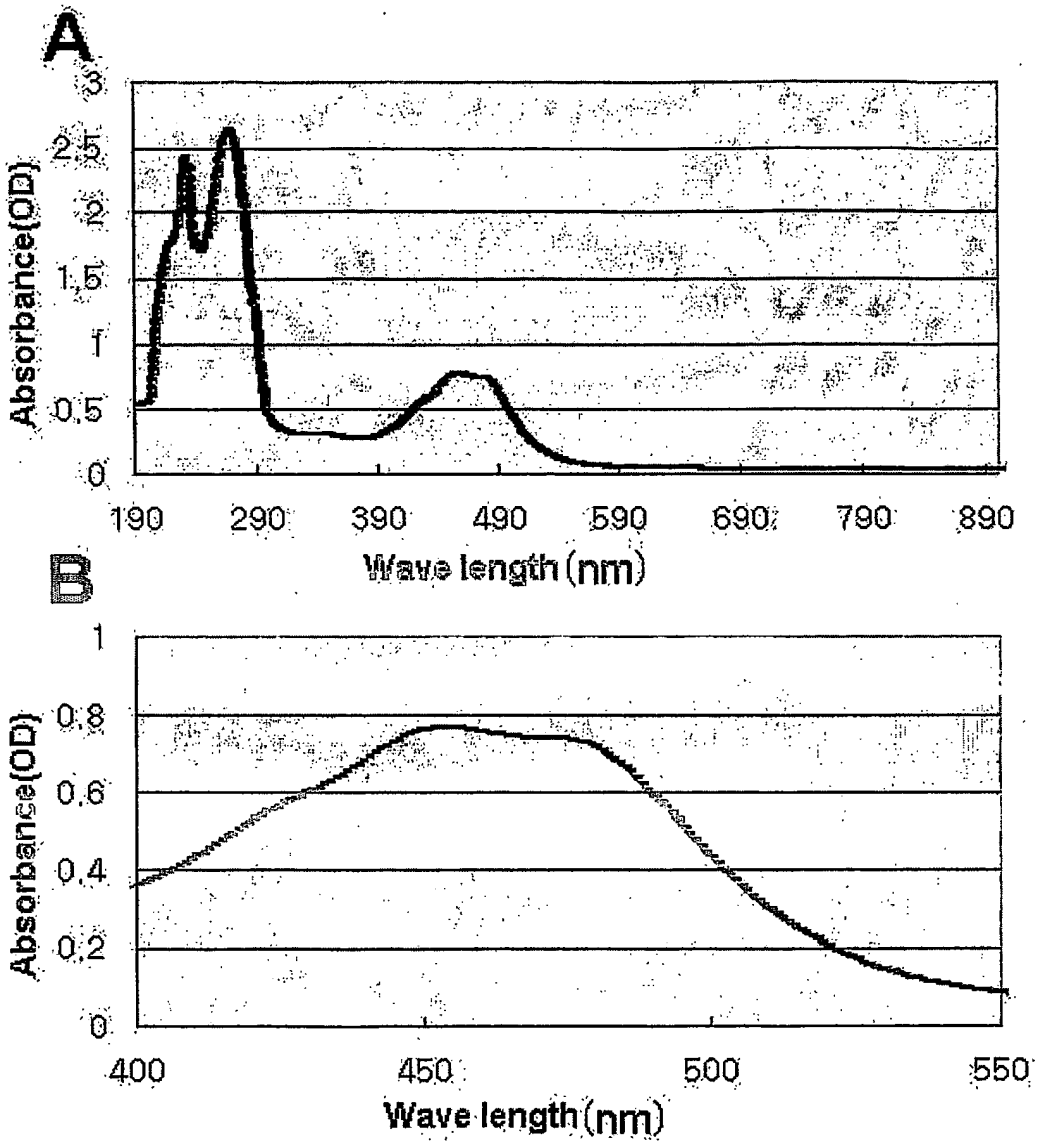

A
Absorbance(OD)
3
2.5
2
1.5
1
0.5
0
190
290
390
490
590
690
790
890
Wave length (nm)
B
Absorbance(OD)
1
0.8
0.6
0.4
0.2
0
400
450
500
550
Wave length (nm)

FIG. 8

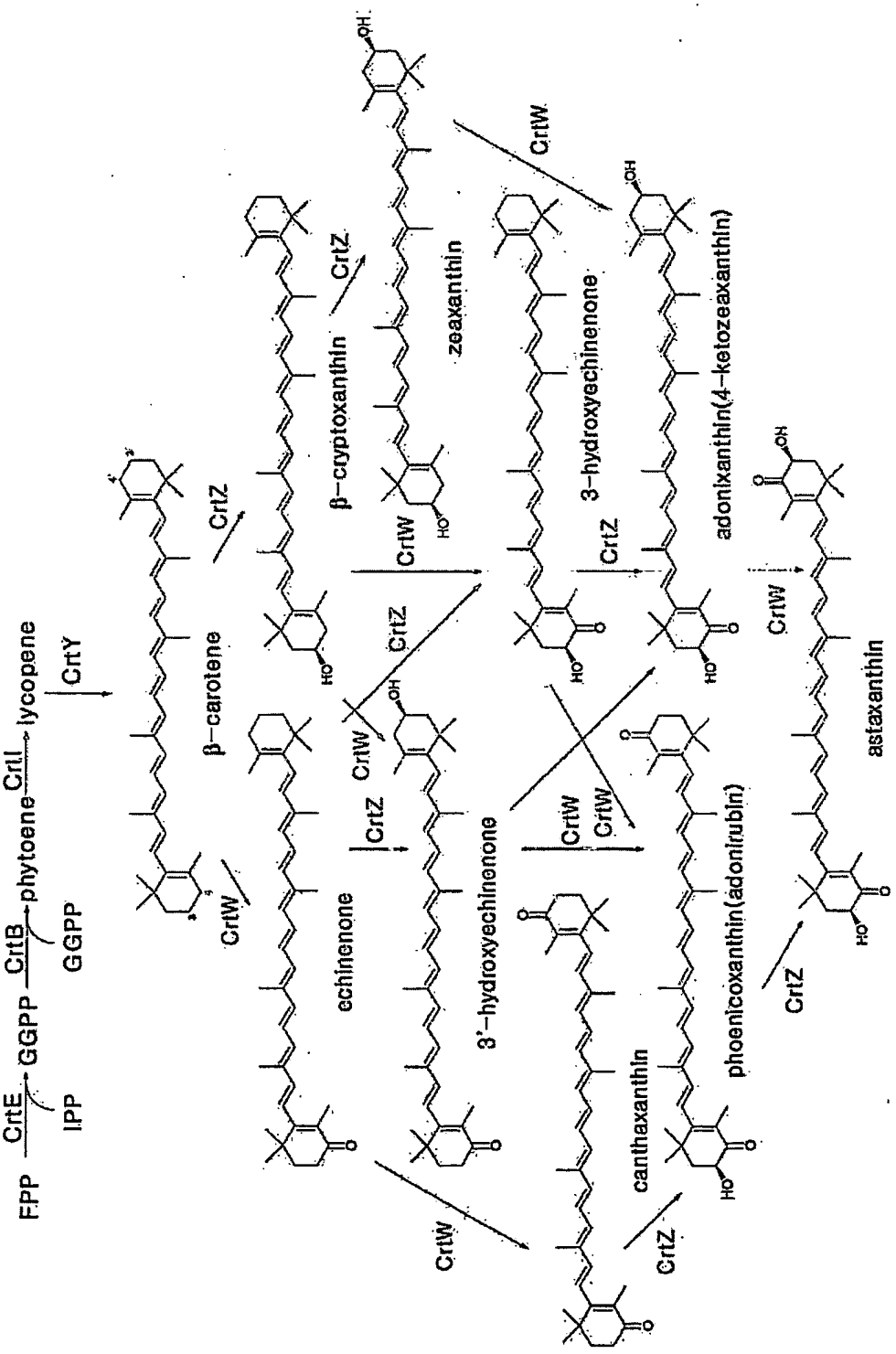

FPP
CrtE
GGPP
IPP
CrtB
phytoene
CrtI
lycopene
CrtY
β-carotene
CrtZ
CrtW
β-cryptoxanthin
zeaxanthin
echinenone
3'-hydroxyechinenone
3-hydroxyechinenone
adonixanthin(4-ketozeaxanthin)
canthaxanthin
phoenicoxanthin(adonirubin)
astaxanthin

FIG. 9

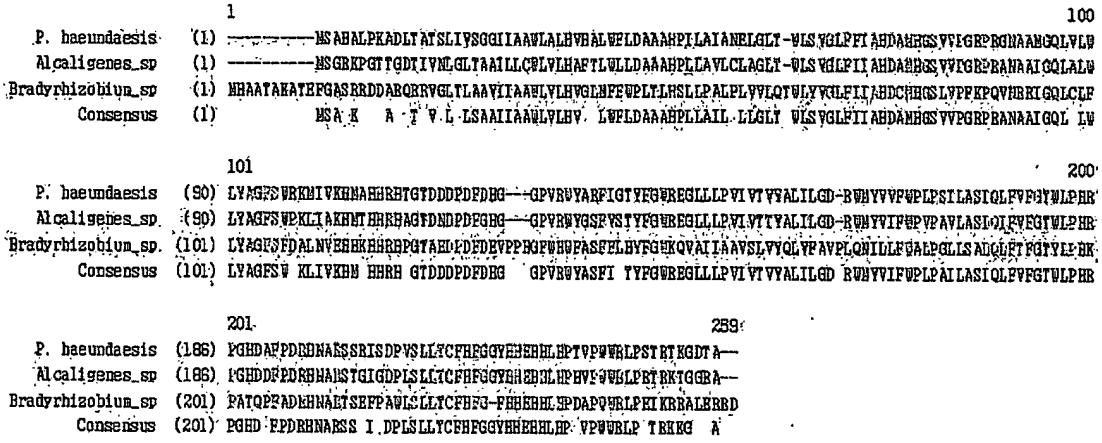

1 100
P. haeundaesis (1) ----------MSAHALPKADLTATSLIYSGGIIAAWLALHVHALWFLDAAAHPILAIANELGLT-WLSVGLPFIAHDAMHGSVVIGRPRGNAAMGQLVLW
Alcaligenes_sp (1) ----------MSGRKPGTTQDTIVMLGLTAAILLCGLVLHAFTLWLLDAAAHPLLAVLCLAGLT-WLSVGLPIIAHDAMHGSVVIGRPRANAAIGQLALW
Bradyrhizobium_sp (1) MHAATAHATEFGASRRDDARQRRVGLILAAVIIAAWLVLHVGLHFFWPLTLRSLLPALPLVVLQTWLVVGLFIIAHDCMHGSLVPFKPQVNRRIGQLCLF
Consensus (1) MSA A T V L LSAAIIAAWLVLHV LWFLDAAAHPLLAIL LLGLT WLSVGLFIIAHDAMHGSVVPGRPRANAAIGQL LW
101 200
P. haeundaesis (90) LYAGFSVRKMIVKHNAHHRHTGTDDDPDFDEG---GPVRWYARFIGTYFGWREGLLLPVIVTVYALILGD-RWHYVVFWPLPSILASIQLFVFGTWLPHR
Alcaligenes_sp (90) LYAGFSVPKLIAKHNTHHRHAGTDNDPDFGHG---GPVRWYGSFVSTYFGWREGLLLPVLVTTYALILGD-RWHYVIFWPVPAVLASLQLFVFGTWLPHR
Bradyrhizobium_sp. (101) LYAGFSFDALNVEEHKHHRHPGTAHDPDFDEVPPHGFWHWPASFELHYFGWKQVAIIAAVSLVYQLVFAVPLQNILLFWALPGLLSALQLFTFGTYLPHK
Consensus (101) LYAGFSW KLIVKHN HHRH GTDDDPDFDEG GPVRWYASFI TYFGWREGLLLPVIVTVYALILGD RWHYVIFWPLPAILASIQLFVFGTWLPHR
201 259
P. haeundaesis (186) PGHDAFPDRHNAASSRISDPVSLLTCFHFGGYHEEHHLHPTVPFWRLPSTRTKGDTA--
Alcaligenes_sp (186) PGHDDFPDRHNAHSTGIGDPISLLTCFHFGGYHHEHHLHPHVPFWRLPETRKTGGRA--
Bradyrhizobium_sp (201) PATQPFADHHNAETSEFPAVLSLLTCFHFG-FHEEHHLSPDAPQWRLPETKRRALERRD
Consensus (201) PGHD FPDRHNARSS I DPLSLLTCFHFGGYHHEHHLHP VPWWRLP TRHKG A

FIG. 10

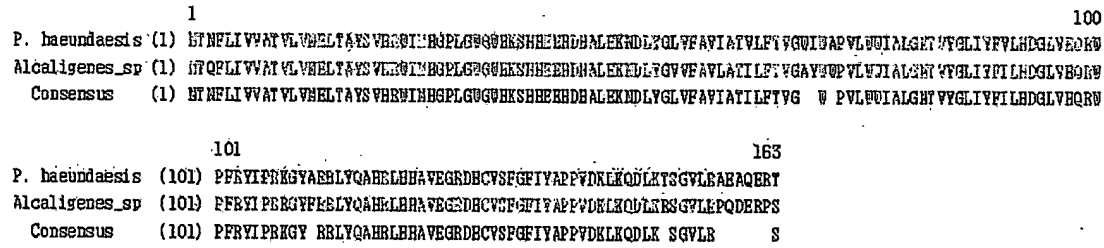

1 100
P. haeundaesis (1) MTNFLIVVATVLVHELTAYSVRWIIHHGPLGWGWHKSHHEHHDHALEKNDLYGLVFAVIATVLFTVGWIHAPVLWWIALGHTVVGLIYFVLHDGLVHQRW
Alcaligenes_sp (1) MTQFLIVVATVLVHELTAYSVLRWIMHGPLGWGWHKSHHEHHDNALEKNDLFGVVFAVLATILFTVGAVHWPVLWWIALGHTVVGLIYFILHDGLVHQRW
Consensus (1) MTNFLIVVATVLVHELTAYSVHRWIHHGPLGWGWHKSHHEHHDHALEKNDLYGLVFAVIATILFTVG W PVLWWIALGHTVVGLIYFILHDGLVHQRW
101 163
P. haeundaesis (101) PFRYIPRKGYAEHLYQAHRLHHAVEGRDHCVSFGFIYAPPVDKLKQDLKTSGVLEAHAQERT
Alcaligenes_sp (101) PFRYIPERGYFRHLYQAHRLHHAVEGRDHCVSFGFIYAPPVDKLKQDLKESGVLEPQDERPS
Consensus (101) PFRYIPRKGY RRLYQAHRLHHAVEGRDHCVSFGFIYAPPVDKLKQDLK SGVLE S

FIG. 11

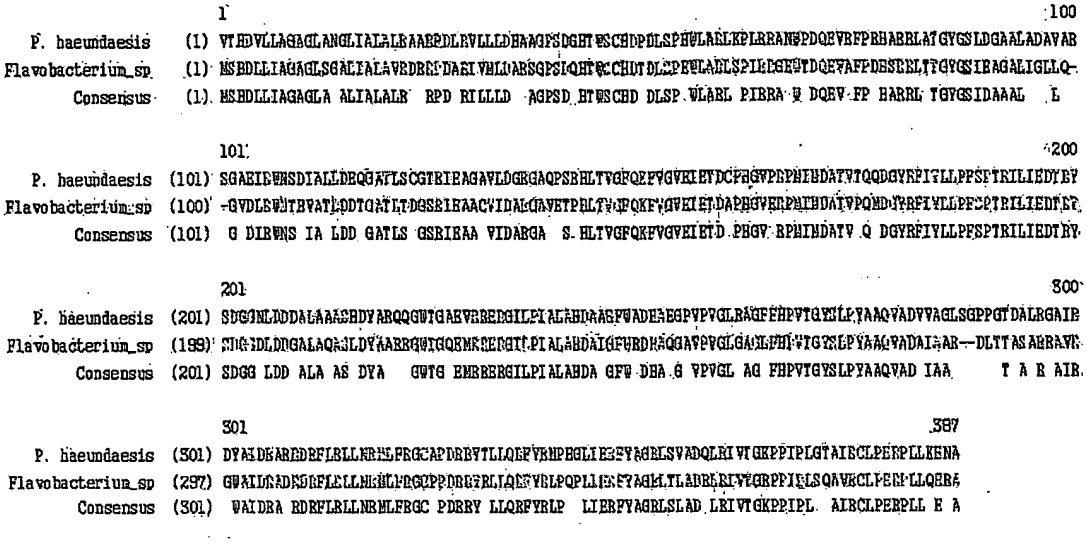


FIG. 12

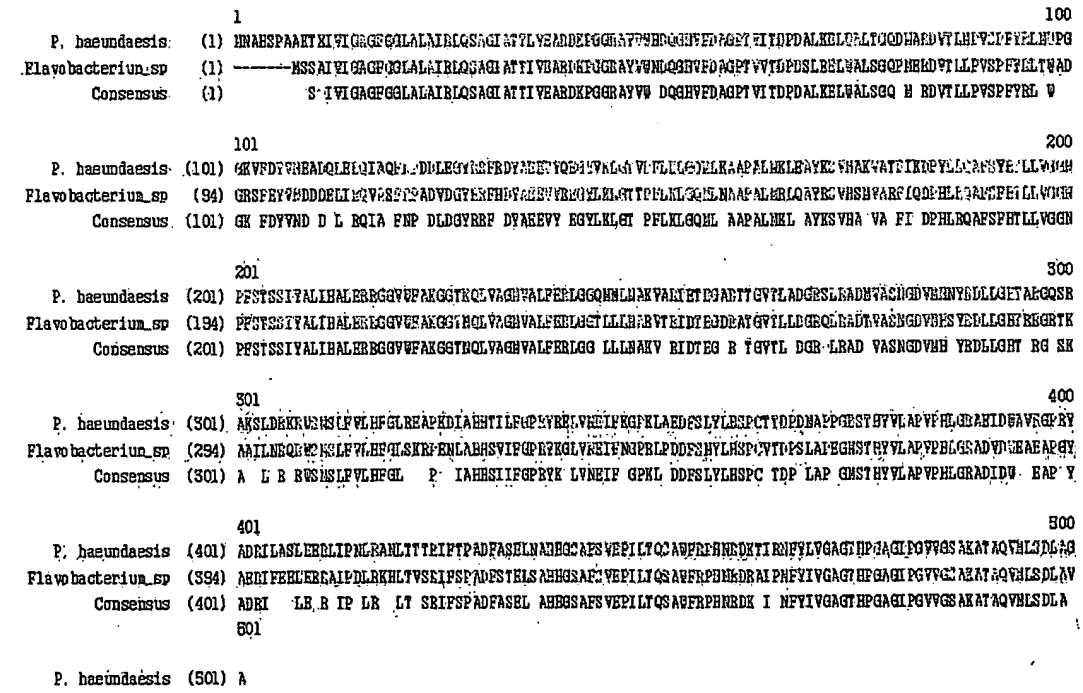


FIG. 13

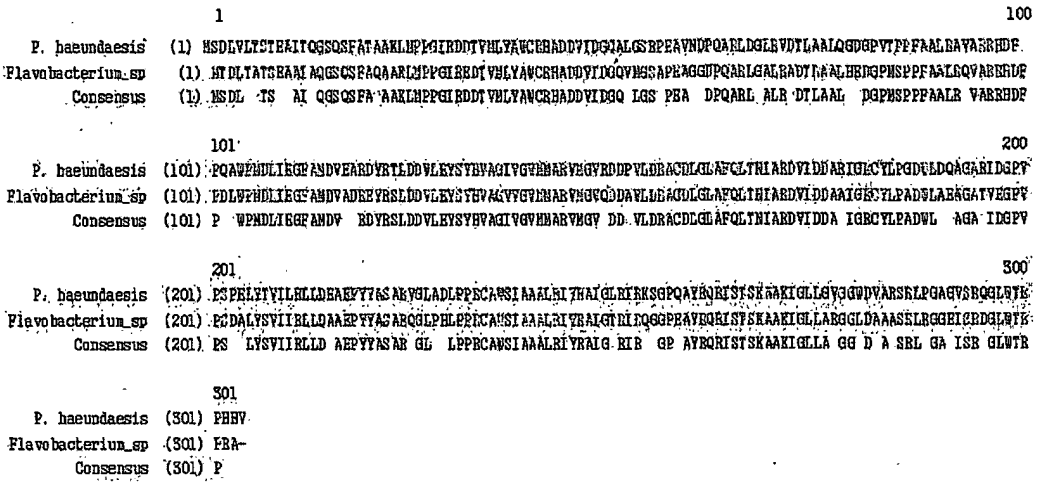


FIG. 14

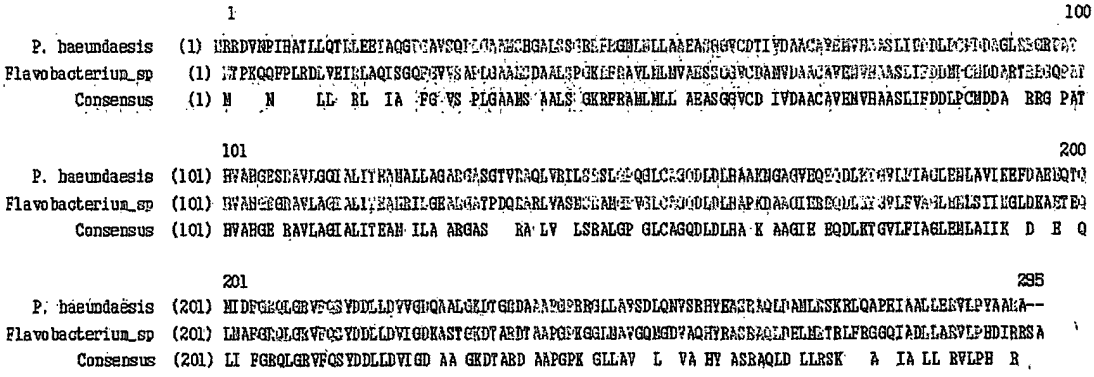
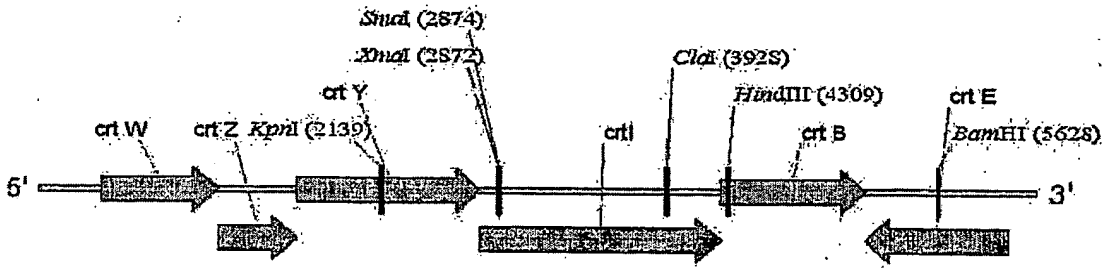


FIG. 15




crt Gene
(6223bp)

FIG. 16

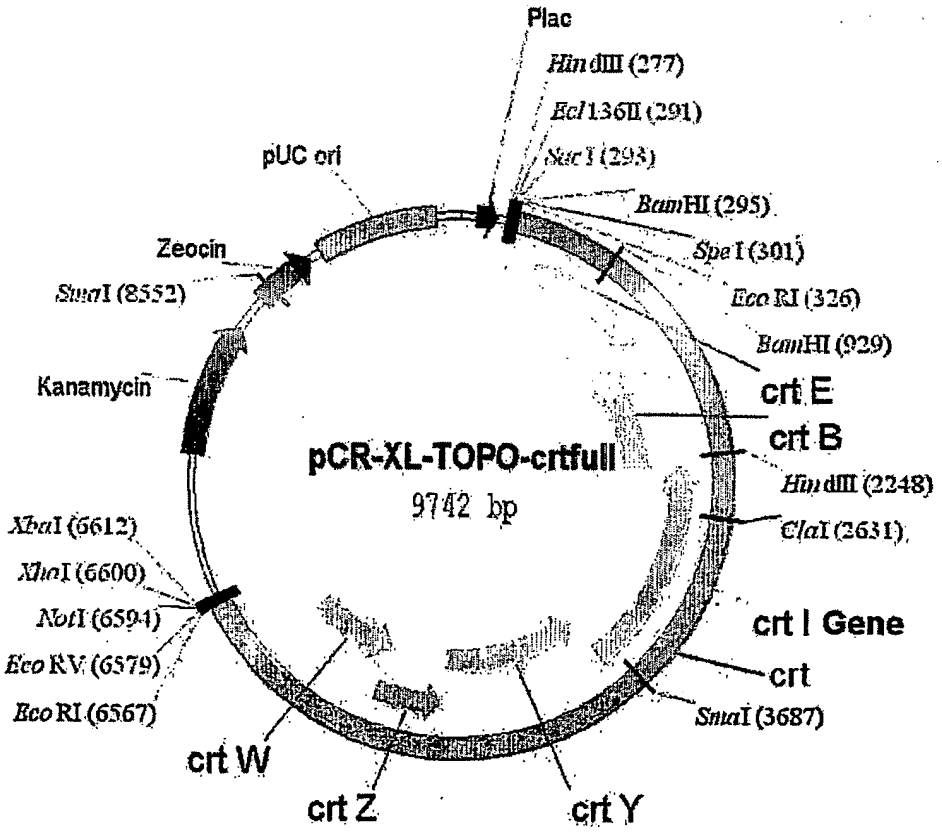
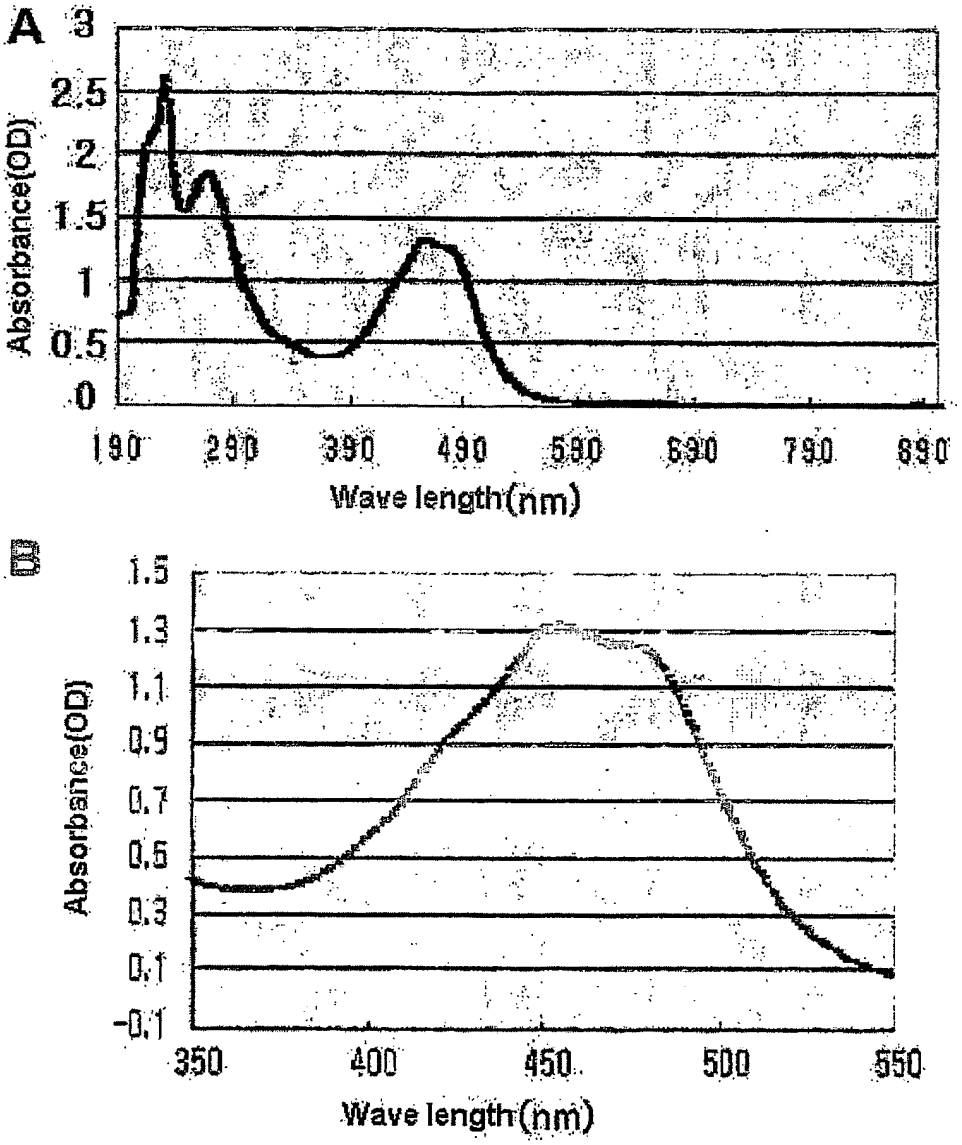


FIG. 17


A
Absorbance(OD)
3
2.5
2
1.5
1
0.5
0
190
290
390
490
590
690
790
890
Wave length(nm)
B
Absorbance(OD)
1.5
1.3
1.1
0.9
0.7
0.5
0.3
0.1
-0.1
350
400
450
500
550
Wave length(nm)

# SEQUENCE LISTING

<110> KIM, Young Tae
LEE, Jae Hyung
ALGENETECH

<120> Gene involved in the biosynthesis of carotenoid and marine microorganism, paracoccus haeundaesis, producing the carotenoid

<130> 4fpo-02-06

<150> KR2003-20222
<151> 2003-03-31

<150> KR2003-20023
<151> 2003-03-31

<160> 18

<170> KopatentIn 1.71

<210> 1
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> forward primer for Paracoccus haeundaesis 16S rDNA

<400> 1
cataagtaat tatggttttg t 21

<210> 2
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> reverse primer for Paracoccus haeundaesis 16S rDNA

<400> 2
cgcttcctta gaaaggag 18

<210> 3
<211> 1454
<212> DNA
<213> Paracoccus haeundaesis

<400> 3
caacttgaga gtttgatcct ggctcagaac gaacgctggc ggcaggctta acacatgcaa 60

gtcgagcgag accttcgggt ctagcggcgg acgggtgagt aacgcgtggg aacgtgccct 120

tctctacgga atagccccgg gaaactggga gtaataccgt atacgccctt tgggggaaag 180

atttatcgga gaaggatcgg cccgcgttgg attaggtagt tggtggggta atggcccacc 240

aagccgacga tccatagctg gtttgagagg atgatcagcc acactgggac tgagacacgg 300

cccagactcc tacgggaggc agcagtgggg aatcttagac aatgggggca accctgatct 360

agccatgccg cgtgagtgat gaaggcctta gggttgtaaa gctctttcag ctgggaagat 420

aatgacggta ccagcagaag aagccccggc taactccgtg ccagcagccg cggtaatacg 480

gagggggcta gcgttgttcg gaattactgg gcgtaaagcg cacgtgggcg gactggaaag 540

tcagaggtga aatcccaggg ctcaaccttg gaactgcctt tgaaactatc agtctggagt 600

tcgagagagg tgagtggaat tccgagtgta gaggtgaaat tcgtagatat tcggaggaac 660

accagtggcg aaggcggctc actggctcga tactgacgct gaggtgcgaa agcgtgggga 720

gcaaacagga ttagataccc tggtagtcca cgccgtaaac gatgaatgcc agacgtcggc 780

aagcatgctt gtcggtgtca cacctaacgg attaagcatt ccgcctgggg agtacggtcg 840

caagattaaa actcaaagga attgacgggg gcccgcacaa gcggtggagc atgtggttta 900

attcgaagca acgcgcagaa ccttaccaac ccttgacatg gcaggaccgc tggagagatt 960

cagctttctc gtaagagacc tgcacacagg tgctgcatgg ctgtcgtcag ctcgtgtcgt 1020

gagatgttcg gttaagtccg gcaacgagcg caacccacgt ccctagttgc cagcattcag 1080

ttgggcactc tatggaaact gccgatgata agtcggagga aggtgtggat gacgtcaagt 1140

tctcatggcc cttacgggtt gggctacaca cgtgctacaa tggtggtgac agtgggttaa 1200

tccccaaaag ccatctcagt tcggattgtc ctctgcaact cgagggcatg aagttggaat 1260

cgctagtaat cgcggaacag catgccgcgg tgaatacgtt cccgggcctt gtacacaccg 1320

cccgtcacac catgggagtt ggttctaccc gacgacgctg cgctaacctt cgggggggcag 1380

gcggccacgg taggatcagc gactggggtg aagtcgtaac aaggtagccg tagggaacc 1440

tgcggctgga tcac 1454

<210> 4
<211> 6223
<212> DNA
<213> crt gene

<400> 4
gttccacgac tggggcatcc ccacgaccgc gtcgctgcgc gccatcgcgc cgatgatggg 60

gccggaccgg gttctggtcg ggtcgggcgg ggtgcgtcac gggctggacg ccgcgcgggc 120
catccgcctc ggcgcggacc tcgtggggca ggcggcccgc gcgctgcccg ccgcgcgcca 180
cagcgccgag gccctgtccg atcacctgtc cgacgtcgtg acccagctgc gcatcgcgat 240
gttctgcacc ggatcgggcg accttgcagc gctgcgctgc gcgcctctgc tggtgccggg 300
gccgggtggc caatggtcgc aagcaacggg gatggaaacc ggcgatgcgg gactgtagtc 360
tgcgcggatc gccggtccgg gggacaagat gagcgcacat gccctgccca aggcagatct 420
gaccgccacc agcctgatcg tctcgggcgg catcatcgcc gcgtggctgg ccctgcatgt 480
gcatgcgctg tggtttctgg acgcggcggc gcatcccatc ctggcgatcg cgaatttcct 540
ggggctgacc tggctgtcgg tcggtctgtt cttcatcgcg catgacgcga tgcacgggtc 600
ggtcgtgccg gggcgtccgc gcggcaatgc ggcgatgggc cagctggtcc tgtggctgta 660
tgccggattt tcgtggcgca agatgatcgt caagcacatg gcccatcacc gccataccgg 720
aaccgacgac gaccccgatt tcgaccatgg cggcccggtc cgctggtacg cgcgcttcat 780
cggcacctat ttcggctggc gcgaggggct gctgctgccc gtcatcgtga cggtctatgc 840
gctgatcctg ggggatcgct ggatgtacgt ggtcttctgg ccgctgccgt cgatcctggc 900
gtcgatccag ctgttcgtgt tcggcacctg gctgccgcac cgccccggcc acgacgcgtt 960
cccggaccgc cataatgcgc ggtcgtcgcg gatcagcgac cccgtgtcgc tgctgacctg 1020
ctttcacttt ggtggttatc atcacgaaca ccacctgcac ccgacggtgc cttggtggcg 1080
cctgcccagc acccgcacca aggggacac cgcatgacca atttcctgat cgtcgtcgcc 1140
accgtgctgg tgatggagtt gacggcctat tccgtccacc gttggatcat gcacggcccc 1200
ctgggctggg gctggcacaa gtcccaccac gaggaacacg accacgcgct ggaaaagaac 1260
gacctgtacg gcctggtctt tgcggtgatc gccacggtgc tgttcacggt gggctggatc 1320
tgggcgccgg tcctgtggtg gatcgctttg ggcatgaccg tctatgggct gatctatttc 1380
gtcctgcatg acgggctggt tcatcagcgc tggccgttcc gctatatccc gcgcaagggc 1440
tatgcccgcc gcctgtatca ggcccaccgc ctgcaccacg cggtcgaggg acgcgaccat 1500
tgcgtcagct tcggcttcat ctatgcgccg ccggtcgaca agctgaagca ggacctgaag 1560
acgtcgggcg tgctgcgggc cgaggcgcag gagcgcacgt gacccatgac gtgctgctgg 1620
caggggcggg ccttgcgaac gggctgatcg ccctggcgct gcgcgcggcg cggcccgacc 1680
tgcgggtgct gctgctggat catgcggcgg gaccgtcaga cggccatacc tggtcctgcc 1740
acgaccccga tctgtcgccg cactggctgg cgcggctgaa gcccctgcgc cgcgccaact 1800

ggcccgacca ggaggtgcgc tttccccgcc atgcccggcg gctggccacc ggttacgggt 1860
cgctggacgg ggcggcgctg gcggatgcgg tggcccggtc gggcgccgag atccgctgga 1920
acagcgacat cgccctgctg gatgaacagg gggcgacgct gtcctgcggc acccggatcg 1980
aggcgggcgc ggtcctggac gggcgcggcg cgcagccgtc gcggcatctg accgtgggtt 2040
tccagaaatt cgtgggcgtc gagatcgaga ccgactgccc ccacggcgtg ccccgcccga 2100
tgatcatgga cgcgaccgtc acccagcagg acgggtaccg attcatctat ctgctgccct 2160
tctctccgac gcgcatcctg atcgaggaca ctcgctattc cgatggcggc aatctggacg 2220
acgacgcgct ggcggcggcg tcccacgact atgcccgcca gcagggctgg accggggccg 2280
aggtccggcg cgaacgcggc atcctgccca ttgcgctggc ccatgacgcg gcgggcttct 2340
gggccgatca cgcggagggg cctgttcccg tgggactgcg cgcggggttc tttcacccgg 2400
tcaccggcta ttcgctgccc tatgcggcgc aggtggcgga cgtggtggcg ggcctgtccg 2460
ggccgcccgg caccgacgcg ctgcgcggcg ccatccgcga ttacgcgatc gaccgggcac 2520
gccgtgaccg ctttctgcgc ctgctgaacc ggatgctgtt ccgcggctgc gcgcccgacc 2580
ggcgctatac cctgctgcag cggttctacc gcatgccgca tggactgatc gaacggttct 2640
atgccggccg gctgagcgtg gcggatcagc tgcgcatcgt gaccggcaag cctcccattc 2700
cccttggcac ggccatccgc tgcctgcccg aacgtccgct gctgaaggaa aacgcatgaa 2760
cgcccattcg cccgcggcca agaccgccat cgtgatcggc gcaggctttg gcgggctggc 2820
cctggccatc cgcctgcagt ccgcgggcat cgccaccacc ctggtcgagg cccgggacaa 2880
gcccggcggg cgcgcctatg tctggcacga tcagggccat gtcttcgacg cgggcccgac 2940
cgtcatcacc gaccccgatg cgctcaagga gctgtgggcg ctgaccgggc aggacatggc 3000
gcgcgacgtg acgctgatgc cggtgtcgcc cttctatcga ctgatgtggc cgggcgggaa 3060
ggtcttcgat tacgtgaacg aggccgatca gctggagcgc cagatcgccc agttcaaccc 3120
ggacgacctg gaaggatacc gccgcttccg tgattacgcg gaggaggtgt atcaggaggg 3180
ctacgtcaag ctgggcaccg tgcccttcct caagctgggc cagatgctca aggccgcgcc 3240
cgcgctgatg aagctggagg cctataagtc cgtccatgcc aaggtcgcga ccttcatcaa 3300
ggacccctat ctgcggcagg cgttttcgta tcacacgctg ctggtgggcg ggaatccctt 3360
ctcgaccagc tcgatctatg cgctgatcca cgcgctggag cggcgcggcg gggtctggtt 3420
cgccaagggc ggcaccaacc agctggtcgc gggcatggtc gcgctgttcg aacggcttgg 3480

cggccagatg atgctgaacg ccaaggtcgc ccggatcgag accgagggcg cgcggaccac 3540

gggcgtcacc ctggcggacg ggcggtcttt aagggccgac atggtcgcca gcaacggcga 3600

cgtcatgcac aactatcgcg acctgctggg ccacacggcc cgcgggcaga gccgcgcgaa 3660

atcgctggac cgcaagcgct ggtccatgtc gttgttcgtg ctgcatttcg gtctgcgcga 3720

ggcgcccaag gacatcgcgc atcacaccat cctgttcggc ccccgctaca gggagctggt 3780

caacgagatc ttcaagggcc cgaagctggc cgaggatttc tcgctgtacc tgcattcgcc 3840

ctgcacgacc gatccggaca tggcgcctcc gggcatgtcc acgcattacg tgctggcccc 3900

cgtgccgcat ctgggccgcg ccgagatcga ttgggcggtc gaggggccgc gctatgccga 3960

ccgcatcctg gcgtccctgg aggagcggct gatcccgaac ctgcgcgcca acctgaccac 4020

gacgcgcatc ttcacgcccg ccgatttcgc cagcgaactg aacgcccatc acggcagcgc 4080

cttctcggtc gagccgatcc tgacgcaatc cgcgtggttc cggccgcaca accgcgacaa 4140

gacgatccgc aacttctatc tggtcggcgc gggcacccat ccgggcgcgg gcattccggg 4200

cgtcgtgggc tcggccaagg ccacggccca ggtgatgctg tccgacctgg cgggcgcatg 4260

agcgatctgg tcctgacctc gaccgaggcg atcacccaag ggtcgcaaag ctttgccacg 4320

gcggccaagc tgatgccgcc gggcatccgc gacgacacgg tgatgctcta tgcctggtgc 4380

cgccacgcgg atgacgtgat cgacggtcag gccctgggca gccgccccga ggcggtgaac 4440

gacccgcagg cgcggctgga cggcctgcgc gtcgacacgc tggcggccct gcagggcgac 4500

ggtccggtga ccccgccctt tgccgcgctg cgcgcggtgg cgcggcggca tgatttcccg 4560

caggcctggc ccatggacct gatcgaaggc ttcgcgatgg atgtcgaggc gcgcgactat 4620

cgcacgctgg atgacgtgct ggaatattcc tatcacgtcg caggcatcgt cggcgtgatg 4680

atggcccgcg tgatgggcgt gcgcgacgat cctgtcctgg accgcgcctg cgacctgggg 4740

ctggcgttcc agctgaccaa catcgcgcgc gacgtgatcg acgatgcgcg catcgggcgg 4800

tgctatctgc cgggggactg gctggaccag gcgggcgcgc ggatcgacgg gccggtgccg 4860

tcgccggagc tgtacacagt gatcctccgg ctgttggatg aggcggaacc ctattacgcg 4920

tcggcgcggg tgggtctggc ggatctgcca ccgcgctgcg cctggtccat cgccgccgcg 4980

ctacggatct atcgcgccat cgggctgcgc atccgcaaga gcgggccgca ggcctatcgc 5040

cagcggatca gcacgtccaa ggctgccaag atcggcctgc tgggcgtcgg gggctgggat 5100

gtcgcgcgat cacgcctgcc gggggcgggc gtgtcgcggc agggcctctg gacccggccg 5160

catcacgtct aggcgcgcgc ggcgtagggc agaacccgtt ccagcagggc cgcgatttcc 5220

ggagcctgaa ggcgcttgct gcgcagcatc gcgtccagtt gggcgcggct ggcctcgtaa 5280

tgacgggaca cgttctgcag gtctgacacg gccagaaggc cgcgccgcgg gccgggggcc 5340

gcggcatcgc gaccggtatc cttgccaagc gccgcctggt cgcccacgac gtccagcagg 5400

tcgtcatagg actggaacac gcggcccagc tgacggccaa agtcgatcat ctgggtctgc 5460

tcctcggcgt cgaactcctt gatcacggcc agcatctcca gcccggcgat gaacagcacg 5520

ccggtcttca ggtcctgttc ctgttcgacc cccgcgccgt tcttggccgc gtgcaggtcc 5580

aggtcctggc cggcgcacag gccctgcggc cccagggacc gcgacaggat ccgcaccagc 5640

tgcgcccgca ccgtgcccga cgcgccgcgc gcaccggcca gcagggccat tgcctcggtg 5700

atcagggcga tgccgcccag cacggcacgg ctttcgccat gcgccacatg ggtcgcgggc 5760

cggccgcggc gcagcccggc atcgtccatg cagggcaggt cgtcgaagat cagcgatgcg 5820

gcatgcacca tctcgaccgc gcaggcggcg tcgacgatcg tgtcgcagac cccgcccgag 5880

gcctctgccg caagcagcat cagcatgccg cggaaccgcc tgcccgacga cagcgcgcca 5940

tggctcatgg ccgcgccgag cggctgcgac acggcaccga atccctgggc gatctcctca 6000

agtctggtct gcagaagggt ggcgtggatc gggttgacgt ctcgtctcat cagtgccttc 6060

gcgcttgggt tctgacctgg cgggaaggtc aggccggggc ggcaccccgt gacccgtcat 6120

ccaccgtcaa cagtccccat gttggaacgg ttcacgcccg attgcgagcc ttttcgacgg 6180

cgacgcgggg tcgcgcggca atttgtccaa caaggtcagt gga 6223

<210> 5
<211> 729
<212> DNA
<213> crtW gene

<400> 5
atgagcgcac atgccctgcc caaggcagat ctgaccgcca ccagcctgat cgtctcgggc 60

ggcatcatcg ccgcgtggct ggccctgcat gtgcatgcgc tgtggtttct ggacgcggcg 120

gcgcatccca tcctggcgat cgcgaatttc ctggggctga cctggctgtc ggtcggtctg 180

ttcttcatcg cgcatgacgc gatgcacggg tcggtcgtgc cggggcgtcc gcgcggcaat 240

gcggcgatgg gccagctggt cctgtggctg tatgccggat tttcgtggcg caagatgatc 300

gtcaagcaca tggcccatca ccgccatacc ggaaccgacg acgaccccga tttcgaccat 360

ggcggcccgg tccgctggta cgcgcgcttc atcggcacct atttcggctg gcgcgagggg 420

ctgctgctgc ccgtcatcgt gacggtctat gcgctgatcc tgggggatcg ctggatgtac 480

gtggtcttct ggccgctgcc gtcgatcctg gcgtcgatcc agctgttcgt gttcggcacc 540

tggctgccgc accgccccgg ccacgacgcg ttcccggacc gccataatgc gcggtcgtcg 600

cggatcagcg accccgtgtc gctgctgacc tgctttcact ttggtggtta tcatcacgaa 660

caccacctgc acccgacggt gccttggtgg cgcctgccca gcacccgcac caagggggac 720

accgcatga 729

<210> 6
<211> 242
<212> PRT
<213> crtW amino acid

<400> 6

```
Met Ser Ala His Ala Leu Pro Lys Ala Asp Leu Thr Ala Thr Ser Leu
 1               5                  10                  15

Ile Val Ser Gly Gly Ile Ile Ala Ala Trp Leu Ala Leu His Val His
            20                  25                  30

Ala Leu Trp Phe Leu Asp Ala Ala Ala His Pro Ile Leu Ala Ile Ala
        35                  40                  45

Asn Phe Leu Gly Leu Thr Trp Leu Ser Val Gly Leu Phe Phe Ile Ala
    50                  55                  60

His Asp Ala Met His Gly Ser Val Val Pro Gly Arg Pro Arg Gly Asn
65                  70                  75                  80

Ala Ala Met Gly Gln Leu Val Leu Trp Leu Tyr Ala Gly Phe Ser Trp
                85                  90                  95

Arg Lys Met Ile Val Lys His Met Ala His His Arg His Thr Gly Thr
            100                 105                 110

Asp Asp Asp Pro Asp Phe Asp His Gly Gly Pro Val Arg Trp Tyr Ala
        115                 120                 125

Arg Phe Ile Gly Thr Tyr Phe Gly Trp Arg Glu Gly Leu Leu Leu Pro
    130                 135                 140

Val Ile Val Thr Val Tyr Ala Leu Ile Leu Gly Asp Arg Trp Met Tyr
145                 150                 155                 160

Val Val Phe Trp Pro Leu Pro Ser Ile Leu Ala Ser Ile Gln Leu Phe
                165                 170                 175

Val Phe Gly Thr Trp Leu Pro His Arg Pro Gly His Asp Ala Phe Pro
            180                 185                 190

Asp Arg His Asn Ala Arg Ser Ser Arg Ile Ser Asp Pro Val Ser Leu
        195                 200                 205
```

Leu Thr Cys Phe His Phe Gly Gly Tyr His His Glu His His Leu His
210 215 220

Pro Thr Val Pro Trp Trp Arg Leu Pro Ser Thr Arg Thr Lys Gly Asp
225 230 235 240

Thr Ala

<210> 7
<211> 489
<212> DNA
<213> crtZ gene

<400> 7
atgaccaatt tcctgatcgt cgtcgccacc gtgctggtga tggagttgac ggcctattcc 60

gtccaccgtt ggatcatgca cggccccctg ggctggggct ggcacaagtc ccaccacgag 120

gaacacgacc acgcgctgga aaagaacgac ctgtacggcc tggtctttgc ggtgatcgcc 180

acggtgctgt tcacggtggg ctggatctgg gcgccggtcc tgtggtggat cgctttgggc 240

atgaccgtct atgggctgat ctatttcgtc ctgcatgacg ggctggttca tcagcgctgg 300

ccgttccgct atatcccgcg caagggctat gcccgccgcc tgtatcaggc ccaccgcctg 360

caccacgcgg tcgagggacg cgaccattgc gtcagcttcg gcttcatcta tgcgccgccg 420

gtcgacaagc tgaagcagga cctgaagacg tcgggcgtgc tgcgggccga ggcgcaggag 480

cgcacgtga 489

<210> 8
<211> 162
<212> PRT
<213> crtZ amino acid

<400> 8
Met Thr Asn Phe Leu Ile Val Val Ala Thr Val Leu Val Met Glu Leu
1 5 10 15

Thr Ala Tyr Ser Val His Arg Trp Ile Met His Gly Pro Leu Gly Trp
20 25 30

Gly Trp His Lys Ser His His Glu Glu His Asp His Ala Leu Glu Lys
35 40 45

Asn Asp Leu Tyr Gly Leu Val Phe Ala Val Ile Ala Thr Val Leu Phe
50 55 60

Thr Val Gly Trp Ile Trp Ala Pro Val Leu Trp Trp Ile Ala Leu Gly
65 70 75 80

Met Thr Val Tyr Gly Leu Ile Tyr Phe Val Leu His Asp Gly Leu Val

85 90 95

His Gln Arg Trp Pro Phe Arg Tyr Ile Pro Arg Lys Gly Tyr Ala Arg
100 105 110

Arg Leu Tyr Gln Ala His Arg Leu His His Ala Val Glu Gly Arg Asp
115 120 125

His Cys Val Ser Phe Gly Phe Ile Tyr Ala Pro Pro Val Asp Lys Leu
130 135 140

Lys Gln Asp Leu Lys Thr Ser Gly Val Leu Arg Ala Glu Ala Gln Glu
145 150 155 160

Arg Thr

<210> 9
<211> 1161
<212> DNA
<213> crtY gene

<400> 9
gtgacccatg acgtgctgct ggcaggggcg ggccttgcga acgggctgat cgccctggcg 60

ctgcgcgcgg cgcggcccga cctgcgggtg ctgctgctgg atcatgcggc gggaccgtca 120

gacggccata cctggtcctg ccacgacccc gatctgtcgc cgcactggct ggcgcggctg 180

aagcccctgc gccgcgccaa ctggcccgac caggaggtgc gctttccccg ccatgcccgg 240

cggctggcca ccggttacgg gtcgctggac ggggcggcgc tggcggatgc ggtggcccgg 300

tcgggcgccg agatccgctg gaacagcgac atcgccctgc tggatgaaca gggggcgacg 360

ctgtcctgcg gcacccggat cgaggcgggc gcggtcctgg acgggcgcgg cgcgcagccg 420

tcgcggcatc tgaccgtggg tttccagaaa ttcgtgggcg tcgagatcga gaccgactgc 480

ccccacggcg tgccccgccc gatgatcatg gacgcgaccg tcacccagca ggacgggtac 540

cgattcatct atctgctgcc cttctctccg acgcgcatcc tgatcgagga cactcgctat 600

tccgatggcg gcaatctgga cgacgacgcg ctggcggcgg cgtcccacga ctatgcccgc 660

cagcagggct ggaccggggc cgaggtccgg cgcgaacgcg gcatcctgcc cattgcgctg 720

gcccatgacg cggcgggctt ctgggccgat cacgcggagg ggcctgttcc cgtgggactg 780

cgcgcggggt tctttcaccc ggtcaccggc tattcgctgc cctatgcggc gcaggtggcg 840

gacgtggtgg cgggcctgtc cgggccgccc ggcaccgacg cgctgcgcgg cgccatccgc 900

gattacgcga tcgaccgggc acgccgtgac cgctttctgc gcctgctgaa ccggatgctg 960

ttccgcggct gcgcgcccga ccggcgctat accctgctgc agcggttcta ccgcatgccg 1020

catggactga tcgaacggtt ctatgccggc cggctgagcg tggcggatca gctgcgcatc 1080

gtgaccggca agcctcccat tccccttggc acggccatcc gctgcctgcc cgaacgtccc 1140

ctgctgaagg aaaacgcatg a 1161

<210> 10
<211> 386
<212> PRT
<213> crtY amino acid

<400> 10

```
Val Thr His Asp Val Leu Leu Ala Gly Ala Gly Leu Ala Asn Gly Leu
 1               5                  10                  15

Ile Ala Leu Ala Leu Arg Ala Ala Arg Pro Asp Leu Arg Val Leu Leu
            20                  25                  30

Leu Asp His Ala Ala Gly Pro Ser Asp Gly His Thr Trp Ser Cys His
        35                  40                  45

Asp Pro Asp Leu Ser Pro His Trp Leu Ala Arg Leu Lys Pro Leu Arg
    50                  55                  60

Arg Ala Asn Trp Pro Asp Gln Glu Val Arg Phe Pro Arg His Ala Arg
65                  70                  75                  80

Arg Leu Ala Thr Gly Tyr Gly Ser Leu Asp Gly Ala Ala Leu Ala Asp
                85                  90                  95

Ala Val Ala Arg Ser Gly Ala Glu Ile Arg Trp Asn Ser Asp Ile Ala
            100                 105                 110

Leu Leu Asp Glu Gln Gly Ala Thr Leu Ser Cys Gly Thr Arg Ile Glu
        115                 120                 125

Ala Gly Ala Val Leu Asp Gly Arg Gly Ala Gln Pro Ser Arg His Leu
    130                 135                 140

Thr Val Gly Phe Gln Lys Phe Val Gly Val Glu Ile Glu Thr Asp Cys
145                 150                 155                 160

Pro His Gly Val Pro Arg Pro Met Ile Met Asp Ala Thr Val Thr Gln
                165                 170                 175

Gln Asp Gly Tyr Arg Phe Ile Tyr Leu Leu Pro Phe Ser Pro Thr Arg
            180                 185                 190

Ile Leu Ile Glu Asp Thr Arg Tyr Ser Asp Gly Gly Asn Leu Asp Asp
        195                 200                 205

Asp Ala Leu Ala Ala Ala Ser His Asp Tyr Ala Arg Gln Gln Gly Trp
    210                 215                 220

Thr Gly Ala Glu Val Arg Arg Glu Arg Gly Ile Leu Pro Ile Ala Leu
```

225 230 235 240

Ala His Asp Ala Ala Gly Phe Trp Ala Asp His Ala Glu Gly Pro Val
245 250 255

Pro Val Gly Leu Arg Ala Gly Phe Phe His Pro Val Thr Gly Tyr Ser
260 265 270

Leu Pro Tyr Ala Ala Gln Val Ala Asp Val Val Ala Gly Leu Ser Gly
275 280 285

Pro Pro Gly Thr Asp Ala Leu Arg Gly Ala Ile Arg Asp Tyr Ala Ile
290 295 300

Asp Arg Ala Arg Arg Asp Arg Phe Leu Arg Leu Leu Asn Arg Met Leu
305 310 315 320

Phe Arg Gly Cys Ala Pro Asp Arg Arg Tyr Thr Leu Leu Gln Arg Phe
325 330 335

Tyr Arg Met Pro His Gly Leu Ile Glu Arg Phe Tyr Ala Gly Arg Leu
340 345 350

Ser Val Ala Asp Gln Leu Arg Ile Val Thr Gly Lys Pro Pro Ile Pro
355 360 365

Leu Gly Thr Ala Ile Arg Cys Leu Pro Glu Arg Pro Leu Leu Lys Glu
370 375 380

Asn Ala
385

<210> 11
<211> 1506
<212> DNA
<213> crtI gene

<400> 11
atgaacgccc attcgcccgc ggccaagacc gccatcgtga tcggcgcagg ctttggcggg 60

ctggccctgg ccatccgcct gcagtccgcg ggcatcgcca ccaccctggt cgaggcccgg 120

gacaagcccg gcgggcgcgc ctatgtctgg cacgatcagg gccatgtctt cgacgcgggc 180

ccgaccgtca tcaccgaccc cgatgcgctc aaggagctgt gggcgctgac cgggcaggac 240

atggcgcgcg acgtgacgct gatgccggtg tcgcccttct atcgactgat gtggccgggc 300

gggaaggtct tcgattacgt gaacgaggcc gatcagctgg agcgccagat cgcccagttc 360

aacccggacg acctggaagg ataccgccgc ttccgtgatt acgcggagga ggtgtatcag 420

gagggctacg tcaagctggg caccgtgccc ttcctcaagc tgggccagat gctcaaggcc 480

gcgcccgcgc tgatgaagct ggaggcctat aagtccgtcc atgccaaggt cgcgaccttc 540

atcaaggacc cctatctgcg gcaggcgttt tcgtatcaca cgctgctggt gggcgggaat 600

cccttctcga ccagctcgat ctatgcgctg atccacgcgc tggagcggcg cggcggggtc 660

tggttcgcca agggcggcac caaccagctg gtcgcgggca tggtcgcgct gttcgaacgg 720

cttggcggcc agatgatgct gaacgccaag gtcgcccgga tcgagaccga gggcgcgcgg 780

accacgggcg tcaccctggc ggacgggcgg tctttaaggg ccgacatggt cgccagcaac 840

ggcgacgtca tgcacaacta tcgcgacctg ctgggccaca cggcccgcgg gcagagccgc 900

gcgaaatcgc tggaccgcaa gcgctggtcc atgtcgttgt tcgtgctgca tttcggtctg 960

cgcgaggcgc ccaaggacat cgcgcatcac accatcctgt tcggcccccg ctacagggag 1020

ctggtcaacg agatcttcaa gggcccgaag ctggccgagg atttctcgct gtacctgcat 1080

tcgccctgca cgaccgatcc ggacatggcg cctccgggca tgtccacgca ttacgtgctg 1140

gcccccgtgc cgcatctggg ccgcgccgag atcgattggg cggtcgaggg gccgcgctat 1200

gccgaccgca tcctggcgtc cctggaggag cggctgatcc cgaacctgcg cgccaacctg 1260

accacgacgc gcatcttcac gcccgccgat ttcgccagcg aactgaacgc ccatcacggc 1320

agcgccttct cggtcgagcc gatcctgacg caatccgcgt ggttccggcc gcacaaccgc 1380

gacaagacga tccgcaactt ctatctggtc ggcgcgggca cccatccggg cgcgggcatt 1440

ccgggcgtcg tgggctcggc caaggccacg gcccaggtga tgctgtccga cctggcgggc 1500

gcatga 1506

<210> 12
<211> 501
<212> PRT
<213> crtI amino acid

<400> 12

Met Asn Ala His Ser Pro Ala Ala Lys Thr Ala Ile Val Ile Gly Ala
1 5 10 15

Gly Phe Gly Gly Leu Ala Leu Ala Ile Arg Leu Gln Ser Ala Gly Ile
20 25 30

Ala Thr Thr Leu Val Glu Ala Arg Asp Lys Pro Gly Gly Arg Ala Tyr
35 40 45

Val Trp His Asp Gln Gly His Val Phe Asp Ala Gly Pro Thr Val Ile
50 55 60

Thr Asp Pro Asp Ala Leu Lys Glu Leu Trp Ala Leu Thr Gly Gln Asp
65 70 75 80

Met Ala Arg Asp Val Thr Leu Met Pro Val Ser Pro Phe Tyr Arg Leu

```
                85                  90                  95

Met Trp Pro Gly Gly Lys Val Phe Asp Tyr Val Asn Glu Ala Asp Gln
            100                 105                 110

Leu Glu Arg Gln Ile Ala Gln Phe Asn Pro Asp Asp Leu Glu Gly Tyr
        115                 120                 125

Arg Arg Phe Arg Asp Tyr Ala Glu Glu Val Tyr Gln Glu Gly Tyr Val
    130                 135                 140

Lys Leu Gly Thr Val Pro Phe Leu Lys Leu Gly Gln Met Leu Lys Ala
145                 150                 155                 160

Ala Pro Ala Leu Met Lys Leu Glu Ala Tyr Lys Ser Val His Ala Lys
                165                 170                 175

Val Ala Thr Phe Ile Lys Asp Pro Tyr Leu Arg Gln Ala Phe Ser Tyr
            180                 185                 190

His Thr Leu Leu Val Gly Gly Asn Pro Phe Ser Thr Ser Ser Ile Tyr
        195                 200                 205

Ala Leu Ile His Ala Leu Glu Arg Arg Gly Gly Val Trp Phe Ala Lys
    210                 215                 220

Gly Gly Thr Asn Gln Leu Val Ala Gly Met Val Ala Leu Phe Glu Arg
225                 230                 235                 240

Leu Gly Gly Gln Met Met Leu Asn Ala Lys Val Ala Arg Ile Glu Thr
                245                 250                 255

Glu Gly Ala Arg Thr Thr Gly Val Thr Leu Ala Asp Gly Arg Ser Leu
            260                 265                 270

Arg Ala Asp Met Val Ala Ser Asn Gly Asp Val Met His Asn Tyr Arg
        275                 280                 285

Asp Leu Leu Gly His Thr Ala Arg Gly Gln Ser Arg Ala Lys Ser Leu
    290                 295                 300

Asp Arg Lys Arg Trp Ser Met Ser Leu Phe Val Leu His Phe Gly Leu
305                 310                 315                 320

Arg Glu Ala Pro Lys Asp Ile Ala His His Thr Ile Leu Phe Gly Pro
                325                 330                 335

Arg Tyr Arg Glu Leu Val Asn Glu Ile Phe Lys Gly Pro Lys Leu Ala
            340                 345                 350

Glu Asp Phe Ser Leu Tyr Leu His Ser Pro Cys Thr Thr Asp Pro Asp
        355                 360                 365

Met Ala Pro Pro Gly Met Ser Thr His Tyr Val Leu Ala Pro Val Pro
    370                 375                 380

His Leu Gly Arg Ala Glu Ile Asp Trp Ala Val Glu Gly Pro Arg Tyr
```

385 390 395 400

Ala Asp Arg Ile Leu Ala Ser Leu Glu Glu Arg Leu Ile Pro Asn Leu
405 410 415

Arg Ala Asn Leu Thr Thr Thr Arg Ile Phe Thr Pro Ala Asp Phe Ala
420 425 430

Ser Glu Leu Asn Ala His His Gly Ser Ala Phe Ser Val Glu Pro Ile
435 440 445

Leu Thr Gln Ser Ala Trp Phe Arg Pro His Asn Arg Asp Lys Thr Ile
450 455 460

Arg Asn Phe Tyr Leu Val Gly Ala Gly Thr His Pro Gly Ala Gly Ile
465 470 475 480

Pro Gly Val Val Gly Ser Ala Lys Ala Thr Ala Gln Val Met Leu Ser
485 490 495

Asp Leu Ala Gly Ala
500

<210> 13
<211> 915
<212> DNA
<213> crtB gene

<400> 13
atgagcgatc tggtcctgac ctcgaccgag gcgatcaccc aagggtcgca aagctttgcc 60

acggcggcca agctgatgcc gccgggcatc cgcgacgaca cggtgatgct ctatgcctgg 120

tgccgccacg cggatgacgt gatcgacggt caggccctgg gcagccgccc cgaggcggtg 180

aacgacccgc aggcgcggct ggacggcctg cgcgtcgaca cgctggcggc cctgcagggc 240

gacggtccgg tgaccccgcc ctttgccgcg ctgcgcgcgg tggcgcggcg gcatgatttc 300

ccgcaggcct ggcccatgga cctgatcgaa ggcttcgcga tggatgtcga ggcgcgcgac 360

tatcgcacgc tggatgacgt gctggaatat tcctatcacg tcgcaggcat cgtcggcgtg 420

atgatggccc gcgtgatggg cgtgcgcgac gatcctgtcc tggaccgcgc ctgcgacctg 480

gggctggcgt tccagctgac caacatcgcg cgcgacgtga tcgacgatgc gcgcatcggg 540

cggtgctatc tgccggggga ctggctggac caggcgggcg cgcggatcga cgggccggtg 600

ccgtcgccgg agctgtacac agtgatcctc cggctgttgg atgaggcgga accctattac 660

gcgtcggcgc gggtgggtct ggcggatctg ccaccgcgct gcgcctggtc catcgccgcc 720

gcgctacgga tctatcgcgc catcgggctg cgcatccgca agagcgggcc gcaggcctat 780

cgccagcgga tcagcacgtc caaggctgcc aagatcggcc tgctgggcgt cgggggctgg 840

gatgtcgcgc gatcacgcct gccgggggcg ggcgtgtcgc ggcagggcct ctggacccgg 900

ccgcatcacg tctag 915

<210> 14
<211> 304
<212> PRT
<213> crtB amino acid

<400> 14

```
Met Ser Asp Leu Val Leu Thr Ser Thr Glu Ala Ile Thr Gln Gly Ser
 1               5                  10                  15

Gln Ser Phe Ala Thr Ala Ala Lys Leu Met Pro Pro Gly Ile Arg Asp
            20                  25                  30

Asp Thr Val Met Leu Tyr Ala Trp Cys Arg His Ala Asp Asp Val Ile
        35                  40                  45

Asp Gly Gln Ala Leu Gly Ser Arg Pro Glu Ala Val Asn Asp Pro Gln
    50                  55                  60

Ala Arg Leu Asp Gly Leu Arg Val Asp Thr Leu Ala Ala Leu Gln Gly
65                  70                  75                  80

Asp Gly Pro Val Thr Pro Pro Phe Ala Ala Leu Arg Ala Val Ala Arg
                85                  90                  95

Arg His Asp Phe Pro Gln Ala Trp Pro Met Asp Leu Ile Glu Gly Phe
            100                 105                 110

Ala Met Asp Val Glu Ala Arg Asp Tyr Arg Thr Leu Asp Asp Val Leu
        115                 120                 125

Glu Tyr Ser Tyr His Val Ala Gly Ile Val Gly Val Met Met Ala Arg
    130                 135                 140

Val Met Gly Val Arg Asp Asp Pro Val Leu Asp Arg Ala Cys Asp Leu
145                 150                 155                 160

Gly Leu Ala Phe Gln Leu Thr Asn Ile Ala Arg Asp Val Ile Asp Asp
                165                 170                 175

Ala Arg Ile Gly Arg Cys Tyr Leu Pro Gly Asp Trp Leu Asp Gln Ala
            180                 185                 190

Gly Ala Arg Ile Asp Gly Pro Val Pro Ser Pro Glu Leu Tyr Thr Val
        195                 200                 205

Ile Leu Arg Leu Leu Asp Glu Ala Glu Pro Tyr Tyr Ala Ser Ala Arg
    210                 215                 220

Val Gly Leu Ala Asp Leu Pro Pro Arg Cys Ala Trp Ser Ile Ala Ala
225                 230                 235                 240
```

```
Ala Leu Arg Ile Tyr Arg Ala Ile Gly Leu Arg Ile Arg Lys Ser Gly
               245                 250                 255

Pro Gln Ala Tyr Arg Gln Arg Ile Ser Thr Ser Lys Ala Ala Lys Ile
           260                 265                 270

Gly Leu Leu Gly Val Gly Gly Trp Asp Val Ala Arg Ser Arg Leu Pro
       275                 280                 285

Gly Ala Gly Val Ser Arg Gln Gly Leu Trp Thr Arg Pro His His Val
   290                 295                 300
```

<210> 15
<211> 882
<212> DNA
<213> crtE gene

<400> 15

```
atgagacgag acgtcaaccc gatccacgcc acccttctgc agaccagact tgaggagatc   60

gcccagggat tcggtgccgt gtcgcagccg ctcggcgcgg ccatgagcca tggcgcgctg  120

tcgtcgggca ggcggttccg cggcatgctg atgctgcttg cggcagaggc ctcgggcggg  180

gtctgcgaca cgatcgtcga cgccgcctgc gcggtcgaga tggtgcatgc cgcatcgctg  240

atcttcgacg acctgccctg catggacgat gccgggctgc gccgcggccg gcccgcgacc  300

catgtggcgc atggcgaaag ccgtgccgtg ctgggcggca tcgccctgat caccgaggca  360

atggccctgc tggccggtgc gcgcggcgcg tcgggcacgg tgcgggcgca gctggtgcgg  420

atcctgtcgc ggtccctggg gccgcagggc ctgtgcgccg gccaggacct ggacctgcac  480

gcggccaaga acggcgcggg ggtcgaacag gaacaggacc tgaagaccgg cgtgctgttc  540

atcgccgggc tggagatgct ggccgtgatc aaggagttcg acgccgagga gcagacccag  600

atgatcgact ttggccgtca gctgggccgc gtgttccagt cctatgacga cctgctggac  660

gtcgtgggcg accaggcggc gcttggcaag gataccggtc gcgatgccgc ggcccccggc  720

ccgcggcgcg gccttctggc cgtgtcagac ctgcagaacg tgtcccgtca ttacgaggcc  780

agccgcgccc aactggacgc gatgctgcgc agcaagcgcc ttcaggctcc ggaaatcgcg  840

gccctgctgg aacgggttct gccctacgcc gcgcgcgcct ag                      882
```

<210> 16
<211> 293
<212> PRT
<213> crtE amino acid

<400> 16

```
Met Arg Arg Asp Val Asn Pro Ile His Ala Thr Leu Leu Gln Thr Arg
 1               5                  10                  15

Leu Glu Glu Ile Ala Gln Gly Phe Gly Ala Val Ser Gln Pro Leu Gly
            20                  25                  30

Ala Ala Met Ser His Gly Ala Leu Ser Ser Gly Arg Arg Phe Arg Gly
        35                  40                  45

Met Leu Met Leu Leu Ala Ala Glu Ala Ser Gly Gly Val Cys Asp Thr
    50                  55                  60

Ile Val Asp Ala Ala Cys Ala Val Glu Met Val His Ala Ala Ser Leu
65                  70                  75                  80

Ile Phe Asp Asp Leu Pro Cys Met Asp Asp Ala Gly Leu Arg Arg Gly
                85                  90                  95

Arg Pro Ala Thr His Val Ala His Gly Glu Ser Arg Ala Val Leu Gly
            100                 105                 110

Gly Ile Ala Leu Ile Thr Glu Ala Met Ala Leu Leu Ala Gly Ala Arg
        115                 120                 125

Gly Ala Ser Gly Thr Val Arg Ala Gln Leu Val Arg Ile Leu Ser Arg
    130                 135                 140

Ser Leu Gly Pro Gln Gly Leu Cys Ala Gly Gln Asp Leu Asp Leu His
145                 150                 155                 160

Ala Ala Lys Asn Gly Ala Gly Val Glu Gln Glu Gln Asp Leu Lys Thr
                165                 170                 175

Gly Val Leu Phe Ile Ala Gly Leu Glu Met Leu Ala Val Ile Lys Glu
            180                 185                 190

Phe Asp Ala Glu Glu Gln Thr Gln Met Ile Asp Phe Gly Arg Gln Leu
        195                 200                 205

Gly Arg Val Phe Gln Ser Tyr Asp Asp Leu Leu Asp Val Val Gly Asp
    210                 215                 220

Gln Ala Ala Leu Gly Lys Asp Thr Gly Arg Asp Ala Ala Ala Pro Gly
225                 230                 235                 240

Pro Arg Arg Gly Leu Leu Ala Val Ser Asp Leu Gln Asn Val Ser Arg
                245                 250                 255

His Tyr Glu Ala Ser Arg Ala Gln Leu Asp Ala Met Leu Arg Ser Lys
            260                 265                 270

Arg Leu Gln Ala Pro Glu Ile Ala Ala Leu Leu Glu Arg Val Leu Pro
        275                 280                 285

Tyr Ala Ala Arg Ala
    290
```

<210> 17
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> forward primer for crt gene

<400> 17
gttccacgac tggggcatc 19

<210> 18
<211> 28
<212> DNA
<213> Artificial Sequence

<220>
<223> reverse primer for crt gene

<400> 18
tccactgacc ttgttggaca aattgccg 28

# INTERNATIONAL SEARCH REPORT

International application No.
PCT/KR2004/000752

## A. CLASSIFICATION OF SUBJECT MATTER

**IPC7 C12N 1/20**

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC7 C12N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched
Korean patents and applications for inventions since 1975

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
PubMed, NCBI, KIPASS, Delphion, CA, PAJ

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

| Category* | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. |
|---|---|---|
| X | NCBI accession No. D58420 (28 August 2002) | 1,2,3 |
| X | NCBI accession No. Y15112 (15 September 1999) | 1,4-8, 10-17 |
| X | NCBI accession No. BAA09591 (28 August 2002) | 9 |
| X | NCBI accession No. CAB56060 (15 September 1999) | 9 |
| X | NCBI accession No. BAA09593 (28 August 2002) | 9 |
| X | NCBI accession No. CAB56062 (15 September 1999) | 9 |
| X | NCBI accession No. CAB56063 (15 September 1999) | 9 |
| X | NCBI accession No. CAB56064 (15 September 1999) | 9 |

☐ Further documents are listed in the continuation of Box C.

☐ See patent family annex.

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier application or patent but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents,such combination being obvious to a person skilled in the art
"&" document member of the same patent family

| Date of the actual completion of the international search | Date of mailing of the international search report |
|---|---|
| 28 JULY 2004 (28.07.2004) | 29 JULY 2004 (29.07.2004) |


Name and mailing address of the ISA/KR
Korean Intellectual Property Office
920 Dunsan-dong, Seo-gu, Daejeon 302-701,
Republic of Korea
Facsimile No. 82-42-472-7140

Authorized officer
KIM, Jae Hyun
Telephone No. 82-42-481-5591


Form PCT/ISA/210 (second sheet) (January 2004)

# INTERNATIONAL SEARCH REPORT

International application No.
PCT/KR2004/000752

**Box No. I Nucleotide and/or amino acid sequence(s) (Continuation of item1.b of the first sheet)**

1. With regard to any nucleotide and/or amino acid sequence disclosed in the international application and necessary to the claimed invention, the international search was carried out on the basis of :

   a. type of material

   [X] a sequence listing

   [ ] table(s) related to the sequence listing

   b. format of material

   [ ] in written format

   [X] in computer readable form

   c. time of filing/furnishing

   [ ] contained in the international application as filed

   [X] filed together with the international application in computer readable form

   [ ] furnished subsequently to this Authority for the purposes of search

2. [ ] In addition, in the case that more than one version or copy of a sequence listing and/or table relating thereto has been filed or furnished, the required statements that the information in the subsequent or additional copies is identical to that in the application as filed or does not go beyond the application as filed, as appropriate, were furnished.

3. Additional comments: